The effects of social and symbolic cues

From the lab to the real world

Flora Ioannidou

School of Psychology

University of Lincoln

This thesis is submitted for the degree of

Doctor of Philosophy

College of Social Science October 2020

Table of contents

List of figures

List of tables

Abstract

Previous studies of social attention have suggested that another person's gaze direction automatically shifts attention in an observer. This has seemed to be a unique ability for gaze cues, but later this reflexive orientation of attention has also been found in another social (e.g., pointing hands) and symbolic (e.g., arrows) cues. Studies demonstrating these effects, however, have examined the effects of these cues under restricted conditions. Studies have shown cues predominantly at fixation and in isolation, on an otherwise empty background. These two conditions fail to reflect the conditions under which both social and symbolic cues are encountered in the real world. The present work therefore extended previous studies on social attention to more real-world like situations. As a first step, cues were embedded in natural scenes, but kept static, to examine the difference between an empty background and a natural context. Three conditions were tested: (1) free viewing with a single cue per scene; (2) free viewing with multiple, sometimes incompatible cues and (3) a visual search with a single cue per scene. In all three experiments, participants spent more time looking at the two social cues (people gazing or pointing) than the symbolic cue (an arrow sign), suggesting that the presence of another person is enough to draw the observer's attention. While people drew attention, subsequent shifts of attention towards the cued direction were stronger for pointing cues and arrow, both of which aim to signal a direction. The next chapter moved one step further towards real world social attention by moving from static to dynamic cues (video clips). Participants performed one of two task: Free viewing and a memory task. Again, social cues drew observers' attention. As for static cues, the gaze cues did not direct

subsequent attention to the same extent as pointing cues did. The ultimate step towards studying social attention in the real world was made in the final chapter, where participants' eye movements were tracked with a mobile eye tracker while they performed a search task in the presence of social and symbolic cues. The first experiment used cues printed on a piece of paper as stimuli, whereas the second experiment used real people. Possibly due to the search task, the cues were largely ignored and consequently did not elicit a strong cueing effect. From the three cues, pointing gestures most strongly drew observers' attention and produced the strongest cueing effects. Combined, the findings of this thesis suggest that when explored under more realistic environments, gaze cues no longer shift people's attention. Instead other cues more related to direction (e.g., pointing cues) seems to successfully direct the observers' attention. In agreement with previous studies, when in a laboratory setting other people in the scene strongly attract attention. In the real world attention on the cues and cues' effects are limited.

Acknowledgements

First and foremost, i would like to thank my three supervisors, Frouke, Kun and Tim, for their advice and guidance through out the PhD. Thank you for making my dream come true. Also, i would like to extend my gratitude to my fellow PhD students (Filippo, Kristel, Charlotte, Natasha, Rebecca, Nadia, Beth, Sophie, Melanie, Matt) for their support in times of stress. You made these three and a half years more enjoyable. A special thanks to the admin staff and technicians (Zoe, Andria, Ian, Matt, Ferenc, Todd) for their hard work. Of course i would like to thank my friends (Niko, Matina, Greta, Tibi and Heather) and family for listening, counseling and dealing with me. Finally, i would like to dedicate this thesis to my husband Foivos for his love, support and understanding.

Declaration

The author Flora Ioannidou was the primary invesigator and writter in all the work presented in this thesis. Contributions include Perl and Matlab codes in Chapter 2 and 4 (Frouke Hermens), guidance with data interpretation (Frouke Hermens, Kun Guo and Timothy Hodgson) and manuscript corrections (Frouke Hermens, Kun Guo, Timothy Hodgson). Experiment 5 (Chapter 4) was made possible by an Experimental Psychology Society (EPS) grant awarded to Dr Frouke Hermens.

Flora Ioannidou

October 2020

Chapter 1

General Introduction

1.1 Introduction

Already in infancy, humans show a profound interest for other people's eyes. Soon after birth, infants will realize that these two areas in the human face, the eyes, are the source of information which can help them to understand their surroundings (Bayliss, Frischen, Fenske, & Tipper, 2007). Moreover, they will come to find that looking and following another person's eyes will give them a deeper understanding of these people's intentions, emotions and other valuable information for a successful communication (Baron-Cohen, 1995). As the years goes by, the interest in the eyes will persist, improve and become more elaborate.

Research on human gaze behaviour is extensive and is not solely the domain of cognitive psychologists. In fact, the importance of gaze cues has been the subject of scientific interest by developmental psychologists, computer scientists and evolutionary biologists. Although, some of these areas seem not to align (in terms of general subject and methodology), they do agree that humans have the tendency to follow another person's gaze direction (Emery, 2000) . Behavioural studies have found that humans are quick in detecting another person's eye movements and respond to even the slightest change in another person's gaze direction (Anderson, Risko, & Kingstone, 2011). Evolutionary biologists have suggested that the way the human eye has evolved allows for an easier detection of gaze direction by others (Tomasello, Hare, Lehmann, & Call, 2007).

However, the ability to follow directional signals from another person is not restricted to following gaze direction. Other biological (e.g., pointing hands) and non-biological (e.g., arrows) cues can have very similar effects in the observer. Although the understanding of the meaning of these cues is formed later on in development (e.g., Leung & Rheingold, 1981), both types of cues are still frequently encountered and used in day to day interactions. One may, however, argue that biological cues (i.e., provided by another person) may have developed more strongly due to a longer evolution period. An important question in social attention research is therefore whether such biological cues indeed have stronger effects in the observer.

The present thesis investigates the effects of social (gazing and pointing) and symbolic cues (arrows) on observers' attention in more real-world settings than previously investigated while cues are embedded as part of a natural scene and the real-world. As research on gaze cue (mainly) is vast and widespread across disciplines, this review chapter will focus on the literature specific to the research questions of this thesis and will examine what is already known about how people in scenes attract attention and whether cues provided by these people have effects on attention in the observer.

1.2 Orienting of Attention with social and symbolic cues

1.2.1 Orienting of attention and gaze cues

It has been said that the eyes and face are the mirror of people's mental state, such as intentions, thoughts, desires and beliefs (Baron-Cohen, 1995; Frischen, Bayliss, & Tipper, 2007). When an observer sees another person's face and eyes directed at a certain location, a natural response will be to also shift attention in that direction (e.g., Driver et al., 1999; Friesen & Kingstone, 1998). To gain a better understanding of the cognitive mechanisms underlying attention orienting in response to perceived eye direction, many studies have used Friesen and Kingstone (1998) gaze cueing paradigm.

Friesen and Kingstone (1998) were the first to modify Posner's classic spatial cueing task (Posner, 1980), which aims to uncover covert attention (attention shifts not obvious as eye movements). In the paradigm participants see a centrally or peripherally presented cue, followed by a peripherally presented target. Attention is said to be shifted to the position or direction of the cue, if responses to the target at that location are faster than responses to targets located elsewhere.

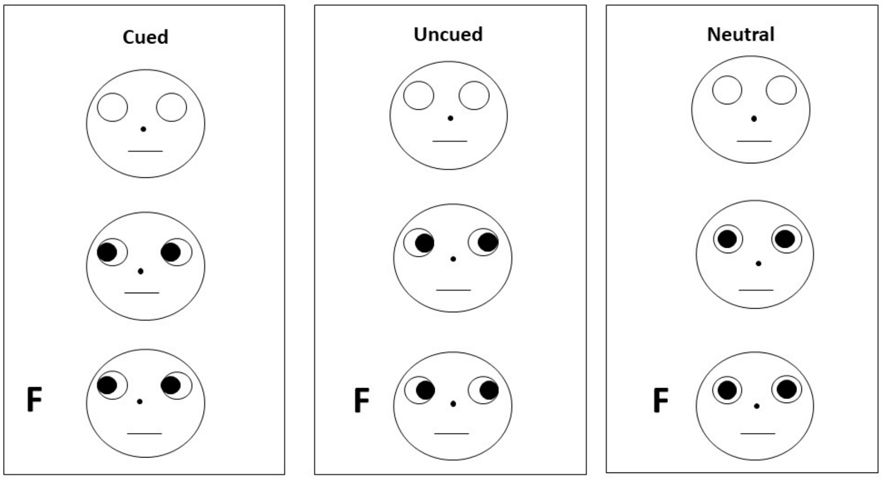

Fig. 1.1 Example of Friesen and Kingstone (1998) the gaze cueing paradigm.

In the gaze cueing version of the paradigm (Figure 1.1), a centrally presented face stimulus is presented with eyes averted to the left or right. Participants are typically informed that the gaze direction does not predict the target's location, and targets are presented equally often in the direction of gaze and the other position. Participants are also told to ignore the cue and to respond to the target as quickly as possible by indicating its direction or its identity (e.g., distinguish between two possible letters). Trials are typically divided into three conditions. In valid trials, the eyes of the centrally presented face are looking towards the future position of the target. In invalid trials, the eyes are looking in the opposite direction. Finally, during neutral trials, the eyes look towards the observer and do not indicate a direction. Typically, valid trials yield faster response times, particularly at short intervals between the onset of the gaze cue and the target (i.e., a short onset asynchrony was around 105ms) and disappear at longer SOAs (around 1,005 ms). These effects of the gaze cue appear to be automatic. They occur even when participants are asked to ignore the cue, and even if the gaze direction is counter-predictive of the target location. Friesen and Kingstone

(1998), with their experiment, started the research field of social attention by providing clear evidence for attention shifts when perceiving another person's gaze shift.

The findings from gaze cueing paradigm were also replicated in children by Hood, Willen, and Driver (1998), who studied attention shifts in children as young as 3 months old. Findings revealed that the gaze cueing effects also occurs in these very young infants, suggesting gaze cues may be present at birth rather than learned by experience.

Further support for gaze cueing was obtained from neuroimaging studies (Hoffman & Haxby, 2000; Hooker et al., 2003; Pelphrey, Singerman, Allison, & McCarthy, 2003; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Sato, Kochiyama, Uono, & Yoshikawa, 2008, 2009). These studies point at the involvement of the posterior superior temporal sulcus (STS) in detecting and processing of eye gaze. The fMRI study by Hoffman and Haxby (2000) found evidence of the involvement of the bilateral intraparietal sulcus and the left posterior STS, by showing stronger activation in the presence of a directional gaze cue than with a non-directional gaze cue. Both areas are related to the superior temporal sulcus in the brain (T. Allison, Puce, & McCarthy, 2000).

Additional evidence for the involvement of the STS region in processing of the direction of gazing eyes was found in human lesion studies (Akiyama et al., 2006; Campbell, Heywood, Cowey, Regard, & Landis, 1990; Eacott, Heywood, Gross, & Cowey, 1993). Findings from these studies show that damage in the STS region greatly impairs the ability to recognize gaze direction. These findings agree with observations by Corbetta (1998), who have shown that the posterior parietal area within the intraparietal sulcus is involved in attentional tasks.

All these evidence has led some researchers (e.g., Hoffman & Haxby, 2000) to suggest that the direction of eye gaze is processed in the superior temporal sulcus and that parietal regions are involved in the automatic shift of attention. However, it is worth mentioning that functional imaging research and electrophysiological studies in humans have found that STS to be activated not only to gaze direction but also to the direction of the head and mouth movement (e.g., T. Allison et al., 2000; George, Driver, & Dolan, 2001; Wicker, Michel, Henaff, & Decety, 1998).

The findings presented so far, strongly suggest that gaze cues affect observers' attention. The question arises whether these responses are caused by the low-level stimulus' properties alone or are related to the biological relevance of the gaze cues. One way to address this question is to compare gaze cues with another cue purely associated with direction.

This has led researchers like Ricciardelli, Bricolo, Aglioti, and Chelazzi (2002) to compare attention shifts in response to gaze cues with those elicited by images of arrows, which do not have biological relevance. Instead of using manual responses to the targets, Ricciardelli et al. (2002) asked participants to make eye movements and to look away from target's direction (i.e., make an anti-saccade). Analysis of the saccade errors showed a higher number of incorrect saccades in the direction of the gaze cue than in the direction of the arrow cue. The anti-saccade task thereby provided strong evidence for reflexive attention in response to gaze cues, but not in response to arrow cues. Although these results seem to provide definite evidence for the importance of biological relevance in cues, subsequent studies failed to provide converging evidence (e.g., potentially because of the different experimental design; Tipples, 2002).

The reflexive orienting of attention for gaze cues has been suggested to be implicitly or explicitly driven by a bottom-up mechanism (e.g., Friesen & Kingstone, 1998; Kingstone, Friesen, & Gazzaniga, 2000). For instance, when using photographs of actual faces, instead of drawings, similar cueing effects are obtained (Driver et al., 1999). Faces were presented as uninformative cues at the center of the display, gaze to the either side of the screen (left or right). On each trial participants were asked to indicate whether a letter (T or L) appeared to either the left or the right of the face. The time between the presentation of the face and the letter varied (the SOA was varied between 100ms, 300ms and 700ms). As in Friesen and Kingstone (1998), valid trials (congruent cuing) showed faster RTs than the invalid trials, but only at the 300 and 700ms SOAs (no difference at the 100ms SOA). In another experiment, the cues were not only unpredictive of the target location but were even counter predictive: The target more often appeared in the opposite direction of that of the gaze cue. Participants could not make use of this information and were still faster when the letter appeared in the direction of gaze, suggestive of reflexive orienting independent of top-down executive control mechanisms, which are sensitive to the predictive nature of a stimulus (e.g., Ristic & Kingstone, 2005).

These results do not mean that there are no top-down influences in gaze cueing (Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002; Dolan et al., 1997). To explore the influence of top-down mechanism on gaze following, Ristic and Kingstone (2005) used a Posner gaze cueing paradigm and presented an ambiguous stimulus at the center of the display. This stimulus could be perceived either as a person with a funny looking hat or as peculiar car. An important feature of this stimulus was that it included a pair of well-defined black border circles, each enclosing a black coloured filled circle. Circles could be identified either as the eyes belonging to the person wearing the hat or as the wheels of the car. One third of the participants were presented with a schematic face similar to the one used in the original gaze cueing paradigm (Friesen & Kingstone, 1998). In all conditions (face and ambiguous stimuli) participants were instructed to ignore the stimuli, as they did not help with target localization and to finish the task as fast as possible. Cueing effect observed by the reaction times on the face and person in a hat stimulus showed the standard gaze cueing behaviour. Both conditions demonstrated slower responses for incongruent to target's location gaze cues. However, these results were different for subjects under the car condition, demonstrating no cost on RTs for incongruent car's wheels.

These evidence provide a strong argument that the effect gaze cues have on observers' attention is not related only to bottom-up mechanisms driven by gaze cues' physical characteristics. Instead top-down mechanisms also contribute in understanding what another person gaze represents. Moreover, the study conducted by Ristic and Kingstone (2005) also highlights one important argument about what is considered as a cue. It can be argued that the results from the person with the hat condition, are not related to the position of the circles representing a biological cue like gazing. Instead, the position of the black circles served as a spatial cue. On the other hand, when on the car condition, wheels did not necessarily depict a cue, as wheel's position did not imply car's direction. Nevertheless, whether incongruent effect caused by the person with a hat condition is a product from its properties, as a basic spatial cue or a distracting social cue, it is with no doubt that gaze cues' visual features are not sufficient enough to lead participant's gaze following effect. In order for the latter to take place, the observer has to identify the presented stimulus as a cue.

The finding that the same stimulus can have two different effects, depending on how the stimulus is interpreted, has led researchers to explore whether gaze cueing involves more than one attentional stream. One such study was by Downing, Dodds, and Bray (2004) who also applied a gaze cueing task. Instead of a gaze cue these researchers used a biological cue, but not commonly used to indicate a direction: Namely a face with a tongue extending to the left or right. These tongue stimuli showed the same kind of cueing as gaze cues, suggesting that any type of directional cue can induce gaze shifts in the observer, but only when such cues are sufficiently predictive of the target location. Gaze cues, in contrast, shift attention even when they are not predictive to target's location suggesting that there are levels of biological cues' effectiveness, with gaze cues providing stronger cueing than tongue cues. One way of interpreting such findings is saying that gaze cues lead to reflexive, exogenous attention shifts, whereas other cues such as tongues and arrows, lead to voluntary, endogenous attention shifts.

The results by Driver et al. (1999) can also be understood in terms of these two attention mechanisms (exogenous and endogenous). At shorter SOAs the reflexive mechanism is at work, whereas at longer SOAs, the voluntary mechanism takes over. The idea of two types of attention is also in line with the results by Kuhn and Kingstone (2009), who asked participants to make eye movements in response to a centrally presented gaze or arrow cues. In line with Driver et al. (1999), Kuhn and Kingstone (2009) found a positive cueing effect for cues with a 50% reliability and for all SOAs tested.

To examine the effect of the reliability of the cue, Hill et al. (2010) systematically varied the reliability of the gaze cue between 100%, 0% or 50%. At short SOAs, response times for 0% reliable gaze cues were significantly slower than those for 100% reliable cues. For cues with 50% reliability congruent gaze cues led to faster response times than incongruent cues, independent of the SOA between target and cue. These results suggest that both reflexive and voluntary attention shifts are at play in gaze cueing. Shorts SOAs combined with low reliability elicit automatic and short-lived attention shifts. Unreliable cues at longer SOAs result in an adaptive cueing effect, suggestive of slower top-down inhibition entering the process.

The experiments so far have focused on eye gaze cueing, meaning that the eyes of the other person need to be visible (e.g., Driver et al., 1999; Friesen & Kingstone, 1998). In the standard gaze cueing paradigm this is easily achieved, but much less common under real-world conditions, where it is much more likely to see a person's head turned without seeing the direction of the eyes within the head. In real-world viewing, the face is also not commonly immediately viewed at fixation, and therefore it is important to investigate whether eye gaze cueing is still occurring when the face is initially viewed from peripheral vision.

In six experiments, Burton, Bindemann, Langton, Schweinberger, and Jenkins (2009) explored the effects of peripherally presented cues. They presented the cues randomly above or below fixation for a short amount of time, to ensure participants could not make an eye movement towards the cue (the time to start an eye movement is around 200ms). The experiments showed that when the gaze cues were presented away from fixation, their cueing effect was weak. This was independent of whether the eyes were presented within a face, or in isolation. Cues that did strongly cue attention were a pointing hand and a rotated head.

Similar results were obtained by Hermens (2017), who used a visual search task, and Hermens, Bindemann, and Burton (2015), who presented the cues for longer and allowed eye movements towards the cues. These results suggest that when eyes are (initially) presented away from fixation they do not strongly cue visual attention, possibly due to visual crowding. This means that in the real-world observers might look for different cues, and head and body orientation may be the ones to look for next.

The influence of head orientation has been studied in several studies. For instance, in neurophysiological studies by Perrett, Hietanen, Oram, and Benson (1992) and Perrett, Oram, Hietanen, and Benson (1994), cells from macaque STS suggested the existence of a hierarchy of coding head and gaze orientation. According to Perrett et al. (1992; 1994), at the top of the hierarchy of social attention cues are the directional gazing, followed by the head orientation and then the direction of the body (Perrett et al., 1985). Moreover, both studies by Perrett et al. (1992; 1994) have also suggested that direction of attention does not exclusively depend on the information conveyed by the eyes. Instead, information from the eyes, head and body are integrated to provide the best possible estimation of where another person is attending. These cues may interact, as behavioural studies (Anstis, Mayhew, & Morley, 1969; Cline, 1967; J. J. Gibson & Pick, 1963) have suggested that head orientation affects the perception of gaze direction.

Another study that examined the influence of head direction, but now for upward and downward looking actors, was performed by Langton, Watt, and Bruce (2000) and Langton. (2000). They demonstrated that when head and gaze direction are incongruent, judging these cues' direction impaired. These findings imply, in contrast to Perrett et al. (1992) hierarchical model, head orientation plays a more important role in selective social attention than gaze direction. Further evidence for such an interpretation was obtained by Langton and Bruce

(1999) who presented images of heads looking in different directions (left and right, up and down). In these images, actor's eyes were not easy to distinguish, and they were always in the same direction as the head. Reaction times were fastest when the head was directed towards the target position.

The studies so far have always used stimuli which head and gaze direction were the same. Hietanen (2002), manipulated the orientation of both head and eyes (e.g., head rotated 30° and eyes looking at the observer) and found a smaller cueing effect than (Langton & Bruce, 1999). A possible reason is that in the stimuli, the eyes were looking directly at the observer, establishing joint attention and biasing attention accordingly. These contradictive results do not reduce the importance of head direction in social attention. In fact, they support the hierarchical model of social attention Perrett et al. (1992; 1994) proposed.

Besides head and gaze direction, body direction also provides a cue of attention. Using a spatial cueing task, Hietanen (2002) presented head-body cues before targets' appearance. Cues were provided under different combinations (1) both head and body looking straight ahead; (2) body position forward and head and eyes rotated sideways, and (3) body, head and eyes rotated to the same direction to the left or right. As in Hietanen (1999), where heads and gaze of different orientation were only used, the study also found evidence for the integration of information from different sources (head, body, and gaze).

More evidence for the importance of head orientation and body posture was also reported by Smilek, Birmingham, Cameron, Bischof, and Kingstone (2006), who collected self-reports and eye movement measurements. While eye movements did not suggest participants used head and body posture, participants' self-reports did point at the use of these cues. As a possible reason for the difference between the two measures, Smilek et al. (2006) suggested that eye movement measurements may fail to detect what subjects report.

In a more recent study, Azarian, Buzzell, Esser, Dornstauder, and Peterson (2017) confirmed the self-report evidence from Smilek et al. (2006). In a classic gaze cueing paradigm Azarian et al. (2017) participants moved their eyes to the location of a peripherally presented target. Non-predictive averted body postures were presented prior the onset of the target. Saccades were initiated faster when the body direction was congruent to the targets location, supporting the notion that body posture is capable of eliciting an automatic attentional shift.

Recently, some researchers have started to argue that social attention needs to be studied in more to the real-world conditions (e.g., Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003). Literatures on gaze cues shifting effect mainly include studies conducted in laboratory settings with cues embedded in an empty background and at fixation (where the observer was already looking at). However, in day to day interactions, humans providing the cue are part of an enriched with items environment with their place in the scene varying. Therefore, few researchers (e.g., Kingstone et al., 2003) have argued that the strong cueing effect produced by the eyes may be restricted to the lab and that other cues might become more effective when gaze cues are no longer presented in isolation and on an otherwise empty background.

Studying social attention in the real-world, however, is not without challenges due to the complexity of controlling all possible variables of interest and difficulties in presented many repeated trials to the same participant. As an intermediate step towards studying social attention in the real-world, studies have used photographs of natural scenes, where actors' gazing are compared with other social and non-social cues (e.g., Hermens & Walker, 2015; Zwickel & Võ, 2010). The comparison cue varied across studies. Zwickel and Võ (2010) compared the cueing effect between a person and a loudspeaker. Both cues were presented as part of a computer-generated scene. Gaze cues were considered the head and body direction, rather than the eyes which were impossibly to see. Targets were either cued (the person or the loudspeaker was facing at its direction) or uncued. Results suggested that gaze cues shifted participants' attention to the cued item, for a longer amount of time, sooner and more frequently. The loudspeaker failed to produce similar cueing effects.

While highly controlled, the computer-generated scenes by Zwickel and Võ (2010) may have lacked the rich detail that actual photographs have. Moreover, the control cue (standing loudspeaker) may not have a strong directional weight. Hermens and Walker (2015) therefore repeated the study with photographs of actual scenes using actual people producing the gaze cues (gazing or pointing). To examine the role of the task participants were performing without affecting the presentation duration of the stimulus, participants had to either perform a memory task or a freely view the images. The cueing effect was fairly small for all types of cues, suggesting that cues that are embedded in natural scenes do not automatically attract attention and do not subsequently automatically direct attention. Instead observers seem to be looking at the remainder of the scene.

It is tempting to argue all questions regarding cues embedded in natural scenes have been answered by Zwickel and Võ (2010) and Hermens and Walker (2015). There are, however, remaining questions, for example, whether the size of the stimulus in the image may be important, or the number of competing cues in the scene. All these parameters have been previously explored in studies where the standard gaze cueing paradigm was used (e.g., simultaneously presented cues by Driver et al., 1999). However, it is unclear whether such effects extent to cues embedded in natural scenes.

The focus so far has been on (eye-) gaze cues, but often other social and symbolic cues, such as pointing gestures and arrows, are studied as well. The next two sections therefore will bring forward evidence on these cues, giving a detailed overview how these two cues can equally influence the observer attentional shift.

1.2.2 Orienting of attention and pointing cues

In typical day to day interactions facial expressions, gaze direction and speech are accompanied by hand and arm gestures (e.g., Langton & Bruce, 2000). For example, while in a conversation, speaker can make rhythmical hand movements of different amplitude to place emphasis on what they are saying. The common assumption about gestures and "body language" is that they are an essential communicational tool in the interaction between humans (e.g., Kendon, 1986, 1994; McNeill, 1985, 1987, 1989). An alternative view suggests that gestures are not made for the observer, but instead are an instrument tool in the service of the speaker to maintain a fluent narrative (e.g., Morrel-Samuels & Krauss, 1992; Rimé & Schiaratura, 1991). While gestures are part of the everyday life, the literature on gesture comprehension is rather limited (e.g., Feyereisen & Lannoy, 1991; Langton & Bruce, 2000; Langton, O'Malley, & Bruce, 1996; L. A. Thompson & Massaro, 1986, 1994).

An important kind of gestures is deictic gestures like pointing. For instance, when directing another person towards a certain location in the street the listener can pay attention to the eyes, head and hand. Pointing gestures are special in that they more strongly convey a sense of direction, gestures to describe the shape of an item. However, many researchers (e.g., Efron, 1941, 1972; Ekman & Friesen, 1972; McNeill, 1989) have debated whether pointing gestures are independent or not of speech. It appears that pointing is a gesture falling to both categories: being independent of speech with their own semantic meaning and enhancing the verbal context (Langton et al., 1996). To perceive the direction of pointing gestures, it is not required to see the entire movement. Just seeing the final stage of the pointing gesture suffices to establish the target of the gesture, allowing the use of static images to study the influence of pointing gestures.

Langton et al. (1996) was one of the first to study the effects of deictic gestures in relation to language. In a Stroop interference paradigm (for a review see MacLeod, 1991), deictic gestures (pointing hands) were presented together with spoken words indicating a direction. Participants were asked to either identify the spoken word or to respond to hand's direction. Reaction times revealed a strong interference effect, with pointing cues difficult to be ignored. Langton et al. (1996) set the foundation for exploring the directional properties of gestures like pointing. However, in their study, direction effect was not provided solely by the extended arm and finger (pointing cue). As pointing gestures were provided by an actual actor, orientation could have been influenced by the head and eyes' direction.

Gaze and head orientation strongly bias attention towards the direction they are looking at (for a review see section 1.2.1). To explore if the directional effect of deictic cues reported in Langton et al. (1996) is due to the simultaneous presence of the head, Langton and Bruce (2000) performed a follow-up study in which they presented simultaneously head and pointing cues that were pointing in the same or different directions. The results showed that when the two cues were in the same direction (even if only one of the two cues had to be responded to), response times were faster, suggesting that the direction of the other cue was automatically processed.

Langton and Bruce (2000) were the first to compare the cueing effects of gaze and pointing cues. Moreover, the study also made a strong attempt in presenting the cues in a real-world scenario, by having a single actor produce both cues. In contrast to, for example, Burton et al. (2009) such single actor cues are much more realistic, because in day to day encounters, pointing and gaze cues are rarely seen without the body that is connected to them.

Although Langton and colleagues (Langton & Bruce, 2000; Langton et al., 1996) studies set the foundations exploring the effects pointing cues have to direct people's attention, still studies on this gesture is very limited. Particularly no studies have attempted to explore if pointing cues can show similar cueing effects like those reported to gazing cues. It was not until nine years after Langton and Bruce (2000)'s study that three more studies (Burton et al., 2009; Ivanoff & Saoud, 2009; Sato et al., 2009) compared the effects of pointing and gaze cues. As mentioned earlier, in a six-experiment study, Burton et al. (2009) presented two cues simultaneously and asked subjects to decide on the direction of one of these two cues (cue specified beforehand). Pointing hands produced a congruency effect, meaning that if a task irrelevant pointing hand was in the same direction as the target cue, the direction of the target cue was easier to report. No such congruency effect was found for eye gaze cues. In contrast, head gaze cues had similar interference effects as pointing hands (Burton et al., 2009).

Burton et al. (2009) findings revealed that in peripheral vision pointing hands are a stronger directional cue than gazing eyes but have similar effects to rotated heads. Similar cueing effects were reported in studies where observers could move their eyes to the stimulus (Hermens, 2017; Hermens et al., 2015; Langton & Bruce, 2000). These results suggest that in peripheral vision deictic gestures are stronger than eye-gaze, which may be due to the distinct shape of pointing hands and rotated heads in peripheral vision. It was therefore suggested that a clear outline might be more important than their biological relevance (Hermens et al., 2015).

Such an interpretation is in agreement with crowding effects, where details of stimuli seen in peripheral vision cannot be distinguished if there are other stimuli around the target stimulus (Strasburger, Harvey, & Rentschler, 1991). It may be that pointing gestures have developed for this reason, when it became clear that eye gaze can only be seen effectively if the observer fixates on the eye region. Developmental studies, however, have suggested that responses to pointing gestures develop later than responses to eye gaze, which can already be found at the age of 3-months (e.g., Farroni, Massaccesi, Pividori, & Johnson, 2004; Hood et al., 1998). In contrast (e.g., Lempers, 1976; Leung & Rheingold, 1981) infants start to comprehend pointing gestures are the age of around 9 months old and by the age of 1 years old they also understand pointing to far distant targets (Lempers, 1976). The understanding of pointing gestures is not without limitations. For example, several studies by Butterworth and colleagues (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991) have showed that while following the direction of an adult's gesture, infants younger than 1 year old fixate the first item on their scan path instead. The same studies (Butterworth & Cochran, 1980; Butterworth & Jarrett, 1991) also suggest that infants in their first year of life only attend to items within their visual field. Two longitudinal studies by Morissette, Ricard, and Décarie (1995) and Carpenter, Nagell, Tomasello, Butterworth, and Moore (1998) found that pointing hands only start to shift attention at an average age of 12 months old. From the age of 12 to 18 months old infants start to avoid items that are in front of them and are capable of following an adult's gesture to targets located behind them (Deák, Flom, & Pick, 2000).

The effects of pointing gestures in children also depend on the task. For instance, in an adapted pro-saccade study, Gregory, Hermens, Facey, and Hodgson (2016) showed that saccadic eye movements in children (age 3-10 years old) were strongly influenced by pointing hands and less so by averted eye-gaze. These effects were particularly strong in younger children, which led Gregory et al. (2016) to conclude that although pointing cues' effects are one of the earliest cue children associate with direction, this effect seems to be reduced with increasing age.

In light of these studies, pointing gestures appear to be equally and sometimes more strongly cue attention than gaze cues. One major limitation, however, is that cues were always presented against a blank background (e.g., Burton et al., 2009). A further limitation is that often images of just the head or a hand are used. This is unlike the real-world, where cues are embedded in often complex environments and where the entire person providing the cue can be seen. In an attempt to study such more realistic scenes, Hermens and Walker (2015) presented photographs of indoor and outdoor scenes to participants that contained an actor or an arrow sign. To examine attention, participants' eye movements were tracked. Interestingly, the results were in line with those using isolated hand and head images in that pointing cues affected participants' overt attention more strongly than gaze cues and arrow signs.

While photographs of natural scenes with actors such as those used by Hermens and Walker (2015) are more like the real-world than the isolated images of hands and heads, they may still be rather artificial, because motion information is missing. This is particularly important for pointing cues, which are largely defined by the movement rather than the end state of the movement (it is rather striking that people can infer direction from just the end state). There have been attempts to use dynamic (moving) cues, but the use of such cues has also met some issues (e.g.,Bayliss, di Pellegrino, & Tipper, 2005; Farroni, Johnson, Brockbank, & Simion, 2000). These studies have been restricted to gaze cues (i.e., no pointing gestures) and mostly focus on the development of social attention (i.e., no adult participants).

For instance, one study by Rohlfing, Longo, and Bertenthal (2012) used dynamic deictic cues to explore social attention in infants. Results from this study, suggest that dynamic pointing gestures have a stronger cueing effect in infants (as young as 4.5 months old) than the static versions of these cues. It has, however, been argued that such effects may be largely due to low-level motion perception, and less so to the dynamic aspect of the cue. Moreover, it was revealed that the effect was only found when the hand and finger were pointing in the same direction, suggesting that the shape of the hand was another important factor.

Although gazing is a powerful cue to orient the observers' attention (as discussed in section 1.2.1), the studies presented above showed that another social cue (pointing) can also demonstrate same effects. However, significant cueing effects have also been found by non-biological cues such as arrows. These findings have sparked the debate whether there is something special about social cues (particularly gaze cues) in terms of development and evolution. The remaining part of the introduction will present the literature looking into the differences and similarities between cueing by social and symbolic cues.

1.2.3 Orienting of Attention and Symbolic cues (Arrows)

Attention shifts follow gaze cues have been well established (e.g., Driver et al., 1999; Friesen & Kingstone, 1998). While it was initially suggested such effects were due to the biological nature of the cue, later studies have started to suggest non-biological cues, such as arrows may have the same effects, casting doubt on the relevance of the biological aspect.

The first line of research comparing social and symbolic cues have therefore directly compared gaze and arrow cues within the same cueing paradigm. While most studies seem to suggest biological cues induced reflexive shifts of attention, whether arrows produced the same reflexive attention was unclear. Posner, Nissen, and Ogden (1978) was one of the first to explore the attentional effects of arrows. Participants were presented with a centrally presented arrow and a peripherally presented target and were asked to respond as quickly as they detected the target. From the results it was suggested that arrows, in contrast to sudden peripheral onsets, produce a voluntary, rather than a reflexive shift of attention in the observer. The observer uses the arrow to respond quicker, but the presence of the arrow does not automatically shift peoples' attention.

Expanding Posner et al. (1978) work, Jonides (1981) compared centrally and peripherally presented arrows. Using a paradigm similar to that by Posner, the study asked participants to identify a letter (L or R) among seven distractor letters. Arrows were momentarily presented (flashed) in one of two locations: peripherally near to the location of the target, or centrally, and were presented before the letters appeared on the display. Two different delays between the cue and the letters were used: 90ms to 50ms. Three experiments were performed: (1) comparing the normal version of the task with one in which a memory load task was used, (2) an experiment where the standard task was compared with one in which the validity of the cue was reduced to 12.5% and (3) one in which the percentage of central and peripheral cues was varied. Results from these three experiments suggested that although peripheral cues captured attention even when participants did not expect these cues, in contrast to centrally presented cues, there was no evidence that the arrows produced a reflexive attentional shift.

More than a decade passed since Jonides (1981) study, until two studies started to question whether gaze cues and peripheral onsets were unique in their ability to induce reflexive shifts of attention (Ristic, Friesen, & Kingstone, 2002; Tipples, 2002). Tipples (2002) used a target detection task and in contrast to Jonides (1981) study, arrows were always present on the screen for the duration of the experiment. Under these conditions, Tipples (2002) managed to demonstrate that the reflexive shift of attention is not only restricted to gaze cues, but arrows can show the same effects. In the same year Ristic et al. (2002) addressed the same question, by comparing gaze and arrows cues shifting effect. In three experiments (one testing adults, one testing children, and one a split-brain patient), Ristic et al. (2002) showed that counter-predictive arrow and gaze cues showed very similar cueing effects. Moreover Ristic et al. (2002) study not only showed both types of cues the same strength of cueing, but also the time-course of these effects were indistinguishable.

The two studies by Ristic et al. (2002) and Tipples (2002) sparked a long line of research comparing the cueing effects of arrows and gaze cues (e.g., B. S. Gibson & Bryant, 2005; B. S. Gibson & Kingstone, 2006; Hommel, Pratt, Colzato, & Godijn, 2001; Ristic & Kingstone, 2006; Ristic, Wright, & Kingstone, 2007). These follow-up studies not only looked at the traditional cueing effect, but also examined other evidence initially brought forward for the reflexive nature of attention shifts from gaze cues. For example, the voluntary tendency of observers to shift attention away from the location a gaze cue was directing, is greatly suppressed by the direction of cue, producing a reflexive attentional shift to the cueing location (e.g., Driver et al., 1999; Friesen, Ristic, & Kingstone, 2004). However, this effect was also successfully demonstrated by arrows Tipples (2008), suggesting that gaze and arrow cues can produce the same attentional shifting effects.

While behavioural studies suggested similar cueing effects for gaze and arrow cues, evidence from neuroimaging and brain lesion studies is less conclusive. One brain region that has been associated with gaze processing is the superior temporal cortex (Hoffman & Haxby, 2000; Hooker et al., 2003; Kingstone, Tipper, Ristic, & Ngan, 2004; Puce et al., 1998; Sato et al., 2008). Studies have therefore attempted to establish whether the same area was implicated in the attention shift following arrows. These studies seem to suggest that gaze and arrow cues are processed in different areas (e.g., Akiyama et al., 2006; Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006; Kingstone et al., 2000; Ristic et al., 2002). Others have argued that the reason why two different regions are involved is due to the physical differences in the gaze and arrow stimuli (Tipper, Handy, Giesbrecht, & Kingstone, 2008). Face stimuli, in particular, have been associated with a specific brain area (superior temporal sulcus), and it is was considered to be unlikely that such highly specialized areas would also process arrows. However, Tipper et al. (2008) study provided evidence against this hypothesis. Tipper et al. (2008) used ambiguous cue stimuli that could either be perceived as eyes or an arrow, which allowed for changing the stimulus by instruction to participants. Under both instructions the same activation of the right superior temporal sulcus (STS) was found. The same result was obtained by Sato et al. (2009) who found that the same areas (STS, occipital cortices, inferior frontal gyrus and inferior parietal lobule) were activated for images of actual eyes, hands and arrows.

Similar effects of gaze and arrow stimuli have been found in studies using eye movements to measure shifts of attention, but not in all conditions. One of the first studies exploring if arrow cues can produce an overt shift of attention was by, Ricciardelli et al. (2002) who asked participants to look left or right of fixation, while presenting a centrally presented arrow or gaze cues. Eye movements' recordings showed that congruent cues similarly affected saccades for arrows and gaze cues. In contrast, incongruent gaze cues had stronger effects than incongruent arrow cues, suggesting differential effects of congruent and incongruent gaze and arrow cues.

While the reduced effects of incongruent arrows in Ricciardelli et al. (2002), could be an indication that such cues do not induce reflexive attention, one can also question the type of arrow used. In their study, Ricciardelli et al. (2002) used rather uncommon arrow shapes (e.g., > <), symbols which are also associated with more than one meaning (e.g., suggesting that something has a greater or less value than something else). However, in real life the most common representation of arrows is not just an arrowhead, but as a line with a triangle at the end. Therefore, the effects reported by Ricciardelli et al. (2002) for arrows, might be related to the non-arrow properties caused by the use of these symbols as cues. To address this possible issue, Kuhn and Benson (2007) applied the same saccade task, but now used an actual arrow sign in line with the arrows used in other studies (e.g., Ristic et al., 2002; Tipples, 2002). With these actual arrows, Kuhn and Benson (2007) found that incongruent arrows and gaze cues produced the same number of incorrect saccades, showing that the differences between gaze cues and arrows in Ristic et al. (2002) were due to the arrow shape used. These results were extended by Kuhn and Benson (2007) who also showed similar effects for response latencies in incorrect saccades and by Kuhn and Kingstone (2009).

The studies discussed up to this point all consider adult populations, but a better understanding of the origin of any possible differences between biological (gaze, pointing) and non-biological (arrows) can be obtained by also examining younger participants, who may not have encountered arrows as often as adults. It has been established that human neonates understand that another person's eyes are an important source of information about the environment and as a consequence, following another person's gaze is observed at a very early age (e.g., D'Entremont, Hains, & Muir, 1997). Though difficult to verify experimentally (e.g., argued by Hermens, 2017), it is highly plausible that exposure to arrows occurs later on in life and potentially requires a learning process to fully make use of the information they provide. Studies have therefore tried to determine from which age arrows start to have similar effects as gaze cues (e.g., Gregory et al., 2016; Hermens, 2017).

Testing the effects of social and symbolic cues in very young infants is difficult: Very young infants are unable to verbally respond to stimuli or follow specific response instructions. Typically studies therefore study children from the age of three (e.g., Gregory et al., 2016; Hermens, 2017; Ristic et al., 2002). Interestingly, children from this age tend to show similar cueing by centrally presented arrows and gaze cues (e.g., Barnes, Kaplan, & Vaidya, 2007; Gregory et al., 2016; Kuhn & Tipples, 2010; Ristic et al., 2002), possibly arguing against a long-term learning process for arrows. An alternative explanation, however, is that the learning process for arrows is already complete at the age of three.

Jakobsen, Frick, and Simpson (2013) tested the orienting effects of arrows in infants of nine months old and older children (3 – 4 years old) and found that responses were due to the perceptual properties of the cue, suggesting incomplete learning of the meaning of these cues. Five to six years old children, in contrast, showed behaviour in line with successful learning of the meaning of the cue Jakobsen et al. (2013). The findings by Jakobsen et al. (2013) therefore suggest that the meaning of arrows is acquired, and that this learning occurs in the first years of life.

Collectively, the evidence presented thus far has suggested that strong cueing is not unique to gaze cues, but can also be observed for arrows. Most studies have presented the cues in isolation and at fixation, and an important question therefore is whether the cueing effects extend to more real-life situations where cues are embedded in a background and initially not fixated.

1.3 Why faces and eyes capture attention in the lab and the real-world

Before a cue (either social or symbolic) can start cueing, it needs to be looked at first. Most studies achieve this by presenting cues at fixation, and it has quickly become clear that when this is no longer the case (i.e., cues are presented away from fixation), cueing effects can dramatically change (Burton et al., 2009; Hermens et al., 2015). What is missing from understanding the cueing effect, is the step before cueing: The cue needs to attract the observer's attention (and gaze). When using a free viewing paradigm, actors appear to attract attention more easily than arrows (Birmingham, Bischof, & Kingstone, 2009), but if participants are given other tasks, such gaze behaviour changes and observers' eye movements focus on the task at hand (lab: Buswell, 1935; Yarbus, 1967; real-world : Laidlaw, Foulsham, Kuhn, & Kingstone, 2011).

As early as 1935, Buswell found that an observer's eye movements are strongly influenced by task demands. By using Walter Ufer 'The Solemn People – Taos Indians' Buswell found that participants' eye movements ignored the surrounding landscape, figures' bodies, focusing mainly on people's faces. This study, however, was not as influential as Alfred L. Yarbus', as described in his book "Eye movements and Vision". Using a custom-built eye tracker Yarbus (1967) tracked his own eye movements while he watched various stimuli under different task conditions. His stimuli presented faces of animals, people and sculptures embedded in photographs and paintings. Across images, scan paths suggested that faces and particularly the eyes attract the observer's attention. Yarbus (1967) also observed the strong task effects earlier found by Buswell (1935), but presented these in a very powerful image, directly comparing the scanpaths for the same image across tasks.

While Yarbus (1967) provided evidence that faces in images attract attention, he did not explain why faces draw attention so strongly. Arguably Yarbus (1967) study provided an important insight on gaze behaviour when viewing a face and other people's eye. However, it did not explain why people orient their attention in the presence of a face in a picture. Some low-level mechanisms have been suggested, that they contribute to the understanding of gaze behaviour. For instance, a possible reason may be a photographer bias in combination with a central bias. Photographers tend to place their objects of interest in the center of the display, and people tend to focus on the center of an image (Tatler, 2007). Even in real-world conditions, this central bias seems to persist, suggesting an aversion against looking from the corner of one's eye (Ioannidou, Hermens, & Hodgson, 2016). This central bias tendency seems to be independent of the salient features in the scene. When in an image of a face, eyes are more or less positioned close to the center of the image. Therefore, central bias should lead to more fixation times on this face area. The central bias tendency, however, cannot fully explain the tendency to fixate the eyes inside a face. For instance, Bindemann, Scheepers, and Burton (2009) presented images of heads of different orientations (i.e., from frontal to profile view). A central tendency was found for the first 250ms (middle of image was looked at, no matter where the eyes were), after which attention shifted to the eyes (sometimes away from the center), suggesting that looking at other people's eyes is not influenced by the central fixation bias.

It therefore appears that there are other factors at work that attract attention to the face. These can still be due to low-level mechanisms (Lai, Li, & Wechsler, 2007) aiding face recognition. The tendency to look at the eyes is so strong, that the eyes do not need to be inside the face to attract attention. This was demonstrated by Levy, Foulsham, and Kingstone (2012) who created three types of characters (1) monsters with their eyes not positioned on their face; (2) humanoids with their eyes positioned on their face and (3) humans. For all these characters, eyes were looked early and more often and fixation duration on the eyes did not differ, again suggesting observers seek out the eyes, even if they are not in the center of the display or the object (the face). The tendency to look at the eyes therefore seems less related to a central bias, and more to the information eye gaze can provide in a social interaction to uncover a person's emotions, intentions and thoughts (Baron-Cohen, 1995).

Tendencies to look at faces and the eyes can be overridden by task demands. This was shown by initially was demonstrated by Yarbus (1967) but also by Smilek et al. (2006), who used images of social scenes. Here it is important to distinguish tasks that require information from the eyes and tasks that focus on everything but the people in the scene. For instance, Birmingham, Bischof, and Kingstone (2008b) found that when people were asked to memorise a scene containing two actors, observers spent more time looking at people's eyes in the scene than when freely viewing the scene, suggesting that the face regions are considered as a source of information potentially helping with memorizing scenes better task. Likewise, Itier, Villate, and Ryan (2007) found that when participants were asked to indicate the direction of the eyes or the head in a face stimulus, their eye movements were directed towards the eyes regions for both conditions but they made more initial saccades towards the eyes when judging eye direction. Taken together the outcome from both studies, it seems that the relevance of the eye region determines how strongly this region attracts observers' covert attention.

In most studies, only a single cue is presented on a trial and behaviour compared across trials. These studies will not reveal which cue is stronger when they are placed in direct competition. Competition between cues has been studied for cues embedded in an otherwise empty display (Burton et al., 2009; Hermens et al., 2015; as discussed above), but also for cues embedded in a natural scene (Birmingham et al., 2009). The former two studies showed that gaze cues did not strongly interfere with responses to pointing hands and rotated heads. The latter study showed that people tend to look at people in a scene, even when eyes were not visible, while ignoring arrows.

Apart from task demand and number of cues, social context is another factor that influences whether people in a visual scene attract attention. Birmingham et al. (2009) found that participants fixated actors' faces and eyes for longer when they were seen in a social interaction. In the absence of a social interaction, the time spend looking at other people's faces and eyes was minimized. These results suggest, that in the presence of an interaction, actors' faces and eyes work as an indicator for the observer to understand the context of the scene in which the social event occurs.

All studies so far have used either static stimuli, but no real-world situations. Under laboratory environments, the mere presence of a person (or even just his face), is a powerful element to capture the observers' attention. However, the way eyes have been used and represented in the static images are not the same from the eyes we observe in the real-life interactions. In day to day interactions, it is not only us observing other people's eyes, but the other person can look back at us. Avoiding exploring social attention under real-world scenarios is not without a reason. Repeated trials are much easier to administer when presenting images or videos, and much more complicated when having to create such stimuli in the actual world. If eye movements are recorded, there are additional difficulties of recording them in the real-world (although technology has progressed strongly in recent years), and the amount of work associated with coding fixations in real-world eye movement recordings (because participants can move their heads, no universal regions of interest can be defined across participants).

Although, exploring social attention in the real-world still poses a lot of issues, there are other alternatives to help our understanding of social attention under more realistic scenarios (e.g., the use of dynamic social and symbolic cues). In a recent study, Foulsham, Cheng, Tracy, Henrich, and Kingstone (2010) explored how gaze allocation is influenced by the presence of a social and a dynamic situation. In their study, participants watched four different videos of a group of three people performing a decision-making task. People in the video were assigned to different groups according to their social status. Results suggested that participants spent most of their time looking at the people in the video. Particularly people's face and eyes were the main area of interest compared to the other elements in the scene. More importantly, the results from this study suggest that dynamic social attention can be explored when a more realistic scenario is used. However, when dynamic faces and eyes are not studied under a direct interaction scenario, it can be challenging as it is not clear if it is the motion of the gaze or its biological relevance (e.g., Farroni et al., 2000) that bias people's attention towards them (for more information see Chapter 3, Introduction). Moreover, as in the static stimuli studies, dynamic gaze cues still lack the sense of ecological validity, as participants are still positioned in front of a screen without direct interaction with the environment and cues.

In a pioneer study, Laidlaw et al. (2011) explored actual social interaction in a real-world setting. Participants were asked to wait in a room where either another person was waiting (real person condition) or in a room where a screen showing another person was present (person on a screen condition). In the real-world condition, participants hardly looked at the person (possibly to avoid having to interact), whereas in the person on the screen condition, the person was frequently looked at (in line with the studies examining eye movements towards images. In a follow-up study, Gallup, Chong, and Couzin (2012) showed that people do not only avoid looking at a confederate in a real-life situation, but that they are also reluctant to follow the confederate's gaze direction. These two studies provide clear cues that social attention differs in lab-based settings and real-world conditions. A likely reason that in real-world situations people avoid looking at other people, to minimize the possibility of an unwanted interaction. This indicates that social attention must be studied in the real-world and not just in the lab.

Collectively, the evidence presented in this section highlight the importance faces and eyes have when presented in laboratory setting. However, these effects are without any influence as other parameters (e.g. task) can minimize its effects to capture the observers' attention. When in the real-world faces and eyes are avoided although the number of studies under realistic settings is limited.

1.4 The present thesis

The literature presented in this chapter shows that gaze cues strongly attract and shift observers' attention (e.g., Driver et al., 1999) in laboratory settings. Two real-world studies discussed above (e.g., Gallup, Chong, & Couzin, 2012), have shown that gaze behaviour towards people is demonstrably different in real-world settings, and that social attention must therefore be studied in the real-world. In such real-world settings, cues compete with other elements in the scene and task demands.

For these reasons, my thesis will investigate social attention in more real-world like settings. Instead of directly moving to real-world situations, this thesis will take the approach to take small steps moving from static images to dynamic images, to real-world settings, so that each subsequent step can be compared, and the cause of any difference more easily identified. The thesis will not only study whether people attract attention more strongly than symbolic cues such as arrows (i.e., capture gaze of the observer), but also whether they also cue attention more strongly (i.e., also direct gaze of the observer). It will also explore differences between different social cues: One in which a person has the aim to direct the observer' attention (pointing cues) and one that may just serve to orient attention of the person (gaze cues). Using static images, the effects of cues will be examined (1) in a free viewing task with a single cue (Experiment 1); (2) in a free viewing task with multiple cues in competition and conjunction (Experiment 2) and (3) in a visual search task with informative and uninformative cues (Experiment 3). Cues will be embedded in indoors and outdoors environments. Entire actors rather than just their head or hand will be presented.

The thesis will then move a step closer (Chapter 3, Experiment 4) on how social and symbolic cues are encountered in the real-world (dynamic cues - videos) to examine the effect of motion on social attention. Also, for these stimuli task effects will be studied by comparing free viewing and a memory task. The final step towards the real-world will be made in the final chapter (Chapter 4), where participants will be fitted with a mobile eye tracker and asked to enter rooms with different types of cues. Social and symbolic cues will either be presented as line-drawings on sheets of paper (Experiment 5), or by actual actors in the room (Experiment 6). To give participants a reason to enter the rooms, they will be asked to locate an item in that room.

Together, the studies in this thesis will provide a better understanding of social attention under a more ecologically valid condition. The results will be of importance to understand the nature of social attention, and will help understand situations in which social communication may not be optimal, such as in autism spectrum disorder. The thesis will also develop the tools to study social attention under more valid conditions, and as such the thesis is also expected to provide an important methodological step forward.

Chapter 2

Exploring the effects of social and symbolic cues in natural scenes

2.1 Introduction

Whether gaze behaviour and gestures of others influence attention in an observer more strongly than constructed cues, such as arrows, has long been debated. Previous studies on social attention have suggested that social cues, (e.g., eye gaze and pointing hands), attract observers' interest and capture their attention (e.g.,Birmingham et al., 2009; Fletcher-Watson, Findlay, Leekam, & Benson, 2008). Furthermore, research has shown that social cues also induce shifts of attention in observers (e.g.,Driver et al., 1999; Friesen & Kingstone, 1998). Such effects of gaze have been argued to be automatic, as cues automatically orient observer' attention even when they are counter predictive (e.g.,Friesen et al., 2004).

An important limitation, however, of the majority of these studies is that they have explored the social cues effects presented in isolation (shown on an empty background), at the center of the screen and by using images of cropped (without a body) faces or hands (e.g.,Friesen & Kingstone, 1998; Ristic et al., 2002). The question therefore arises whether social and particularly gaze cues will produce the same effects when they are presented in more naturalistic context (i.e. images of natural scenes).

Recent studies (e.g.,Hermens & Walker, 2015) have veered towards this strategy by exploring the effects of social cues in natural images (i.e. a person pointing or looking at a swing in a playground). In agreement with the classic literature (e.g.,Friesen & Kingstone, 1998) these studies have revealed that when stimuli of actual photographs of people in a scene are used, these people tend to capture observers' attention. This capture of attention also occurs when more than one person was presented in the scene and in instances where actors' faces were always predictive (e.g.,Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010).

One question that still remains is whether people's gazing at a certain location will lead observers to shift their attention to the gazed-at location. Findings from recent studies have suggested that social cues produce strong cueing effect (e.g.,Kuhn, Tatler, & Cole, 2009;

Zwickel & Võ, 2010). For example, in studies where an image of a person was presented in the scene, observers tended to look at it more often and shift their attention to the object that the actor in the scene looked at (e.g.,Hermens & Walker, 2015).

Since the early work of Yarbus (1967) it is known that fixation patterns are greatly influenced by the task participants perform (for more recent examples, see Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010). Such task effects were also found in the context of social attention. For example, in a study conducted by Hermens and Walker (2015) gaze behaviour to social and symbolic cues in natural scenes depended on whether participants viewed images or performed a memory task.

The central focus of social attention has been on gaze cues. However, there are other, possibly stronger social cues that may direct humans' attention and indicate a certain location of an object or a place, such as pointing cues. Studies examining the effects of pointing cues have shown stronger attention shifts when they were presented in the periphery (Burton et al., 2009) and stronger effects of rotated heads and pointing cues than gaze cues (Langton & Bruce, 2000).

Again, most studies on the attention effects of pointing gestures have presented these cues in isolation with nothing else to see in the background. A recent study by Hermens and Walker (2015) has attempted to explore the effects of social (gazing and pointing) and symbolic cues (arrow) when presented in natural scene and away from fixation. Results suggested that both social cues equally captured observers' attention. In terms of steering the observer's attention towards the cued object, pointing hands were the stronger cue.

The automatic shift of orientation caused by social cues direction was thought to be unique only to them (e.g.,Friesen & Kingstone, 1998; Frischen et al., 2007). However, in everyday life situations gaze cues are not the only source of information which can lead to certain location or object. Symbolic cues like arrows are usually the next in line most used cue which has been associated by learning as a directional stimulus. For many years, gaze cues studies have failed to show the same automatic shift of attention with arrows (Jonides, 1981) until the studies of Ristic et al. (2002) and Tipples (2002). In both studies nonpredicitve arrows produced a reflexive orientation effect and showed that the time course of this reflexive attention was similar to eyes.

Following Ristic et al. (2002) and Tipples (2002) studies providing a wealth of evidence (e.g., B. S. Gibson & Bryant, 2005; B. S. Gibson & Kingstone, 2006; Ristic et al., 2007) followed, showing that most of the attentional effects identified with gaze cues can also be found with arrow cues. For example, Tipples (2008) found that arrows can shift an observer's orientation reflexively in a direction other than the observer's intended direction. Before Tipples (2008)'s study such reflexive shifts were considered to be unique to gaze cues. Evidence so far, however, has focused on arrows presented in isolation, except for the studies by Hermens and Walker (2015) and Zwickel and Võ (2010) who have placed cues in natural scenes or computer-generated scenes.

The plethora of these findings suggest the advantage of gazing against other cues (e.g., arrows, pointing hands, directional words) to capture and shift observers' attention. However, these effects are stronger when eyes are centrally presented in the display (e.g., Bayliss & Tipper, 2006b; Driver et al., 1999; Friesen & Kingstone, 1998, 2003; Friesen, Moore, & Kingstone, 2005; Friesen et al., 2004; Frischen et al., 2007). When in the periphery, these effects for gazing cues seems to be minimized and other cues with more distinct shape (e.g., arrow, pointing hand) showing strong cueing effects (e.g., Burton et al., 2009; Hermens, 2017; Hermens et al., 2015).

Moreover, while presented in the periphery and dependent on head's orientation, eyes shape is difficult to be distinguished (e.g., Sun & Balas, 2014). Therefore, head direction seems to be more reliable to judge orientation (e.g., Langton., 2000; Langton et al., 2000). However, due to the limited number of studies it is still unclear if these advantages still persist when 1) gazing cues are embedded in natural scenes and produced by actual people;

2) when regions of interest are better defined and scenes are varied more systematically across cues (compared to Hermens & Walker, 2015); 3) when social and symbolic cues are placed in direct competition or in congruent and 4) whether different cues are still effective when participants are looking for a specific object in the scene (visual search). This study conducted three experiments to examine these questions. Participants viewed photographic images of natural scenes while their eye movements were tracked. To study the effects of social and symbolic cues, images contained different actors that pointed or gazed at a certain location, as well as arrows in the same role.

Experiment 1 aims to confirm the results by Hermens and Walker (2015) with some methodological improvements. More images and better-defined regions of interest (ROIs) will be used. Rectangular regions of interests, previous studies used (e.g., Hermens & Walker, 2015) are not ideal, because they also include substantial parts of the background of the image (especially for pointing cues). Therefore, the present thesis will use a technique inspired by Guo (2012), in which ROIs images of the photographs are created and fixations are superimposed onto these ROIs images. This will allow a much more detailed ROIs and precise analysis. Based on past studies, we expect that gaze cues will attract participants' attention (e.g., Friesen & Kingstone, 1998; Hermens & Walker, 2015) . Pointing cues are expected to lead to most subsequent fixations on the cued target, compared to gaze cues and arrows (Hermens Walker, 2015).

The second experiment examines gaze behaviour when seeing multiple cues. On the basis of past work, people are expected to attract more attention than arrows (Birmingham et al., 2009). A possible issue with this past finding was that arrows were frequently presented alongside a text in the scenes. Moreover, it was unclear whether the different cues led to different target localization. Finally, it is unclear how pointing cues compare to arrows and gaze cues in direct competition.

Similar to Experiment 1 and Hermens and Walker (2015) and Birmingham et al. (2009), images of natural scenes were used. As in Birmingham et al. (2009) study actors and arrows will be placed under competition (e.g., pointing at different directions) or congruent (e.g., pointing at the same direction). In contrast to Birmingham et al. (2009) study, where in some occasions multiple people were presented alongside a single arrow, this study will match only one actor (either pointing or gazing) with an arrow. This manipulation aims to give symbolic cue an equal power in the study. Based on the previous literature in this study we expect that people's faces will attract observers' attention particularly even when they are presented alongside with an arrow or pointing. When exploring the directional effect of the cues, we expect that social cues (particularly pointing) will successfully lead to a certain location in the scene.

Finally, Experiment 3 investigates whether targets in natural scenes can be found more quickly when cued by a social compared to a symbolic cue. Hermens and Walker (2015) already found that a memory task led to different eye movements to social and symbolic cues than free viewing, but their tasks did not involve the cued object. During Experiment 3, participants were presented with images similar to Experiment 1 (containing a single cue) but were now asked to search for a target (presented as a small image before the main image) that was either cued or or not cued (e.g., placed on an opposite location from the direction of the cue). Free viewing task used in Experiment 1 and 2 of this thesis is very commonly used, but has the potential issue that participants overthink the experiment and start wondering if there is a hidden purpose of looking at the images. Moreover, in the real world, free viewing is not a very common activity. In day to day scenarios, people can passively observe a scene happening in front of them (e.g., looking someone asking directions). However, in most of their time, people are actively looking for something (e.g., trying to find an item inside a room). To explore this active behaviour, and to see if this will affect participants' eye movements towards the cues, the visual search task described above was used. Based on research that examined the effects of social and symbolic cues presented in isolation, cued objects are expected to be found more quickly, but no differences are expected between social and symbolic cues (Hermens, 2017; Hermens et al., 2015). Past studies also suggested that gaze-only cues (forward looking face with averted eyes) led to longer search times, but these are not included in the present work (all gaze cues involve a turned head).

2.2 Experiment 1

Experiment 1 aims to replicate the findings of Hermens and Walker (2015) using more precisely defined regions of interest (ROI) and a broader range of stimuli and set the baseline for the two subsequent experiments.

2.2.1 Methods

2.2.1.1 Participants

Thirty participants (27 females), aged 18 to 36 (Mean = 20.70, SD = 4.05), took part in the experiment. Participants were psychology students from the University of Lincoln. They all reported normal or corrected with contact to normal vision. All provided a written consent for the study that was approved by the local ethics committees and received credit courses for their participation.

2.2.1.2 Apparatus

The experiment was conducted in a quiet lab at the University of Lincoln (UK). Eye movements were collected using an Eyelink 1000 (SR Research, Ontario, Canada) desk-mounted eye tracker (at 1,000 Hz, $0.25^{\circ} - 0.5^{\circ}$ average accuracy, 0.01° RMS resolution) and the SR Research Experiment Builder software. This system uses pupil position and corneal reflection to calculate participants' gaze position on the computer screen that they are looking at. Stimuli were presented on a 19 inch Dell flat screen (1280 x 1024 pixels resolution; approximately 26° x 20° of visual angle), placed 80cm away from the participant. Images (Figure 2.1) used in the study were taken with a Samsung Galaxy A3 smart phone camera (camera resolution 1920 x 1080 pixels).

Fig. 2.1 Examples of stimuli used in Experiment 1. Left image (top) showing a person gazing (gazing condition). Right image (top) shows the same person pointing at the same location (pointing condition). Left image (bottom) showing the scene without a cue (no cue condition). Right (bottom) showing an arrow (arrow condition).

2.2.1.3 Stimuli

For the purpose of this study 19 different scenes were photographed showing outdoors and indoors areas of the University of Lincoln (UK) campus and residential areas of Athens, Greece. Each scene was photographed with each of three cues: a) an arrow pointing at a particular object in the scene; b) an actor who was pointing the same object in the scene; c) an actor who was gazing at the same object. Additionally, the scene was photographed with no cue. Ten different actors served as the social cues. Location of the cues was consistent in all conditions and their position was either left (57%) or right (43%). All images were scaled down to the size of 800 x 600 pixels (approximately 16° x 12° of visual angle) with the use of IrfranView software and presented in 1,280 x 1,024 pixels display. Size of the images was scaled down for two reasons: (1) to present fixation point and drift correction around the images location, avoiding biasing participants' attention towards the center of the screen and (2) to prevent Experimental builder re-scale and distort the images as the size of the original images was larger than the size of the screen. Smaller size images should not affect exploring the research questions of this thesis. Previous studies (e.g., Birmingham2009) have presented, in some cases, the images of social and symbolic cues smaller than 800 x 600 pixels and still managed to explore any effect from these two cue categories. Finally, all images were checked after scaling down to ensure that cues and cued objects were clear during the experiment.

2.2.1.4 Design

All participants were presented with all images. This means that the same scene was seen four times by each participant. To avoid that previous views of the scenes influenced the average data, the order of the cues within the scenes was systematically varied across participants by using four blocks of stimuli. These blocks also ensured that each cue was seen equally often in each section of the experiment. This design allows for two analyses: An analysis of all data, making cue a within subjects' factor (with possible repeated exposure effects, but not in the average data), and an analysis of the first block of trials for each participant only, making cue a between subjects' factor (excluding any possible previous exposure effects). The order of the images within each block was randomized for each participant and the order of the block was counterbalanced by means of Latin square.

2.2.1.5 Procedure

At the beginning of the experiment, each participant received written and verbal information from the experimenter and signed the consent form. Then they were instructed to place their chin on the chin rest in order to minimize head movement as much as possible. The eye tracker was calibrated with a nine-point calibration, which was repeated when until the recorded fixations' position was aligned with a three-by-three grid. To avoid fixations at the centre on the image, each trial begun with a drift correction target presented in four different locations at the periphery of the image. This drift correction target also corrected for any small head movements during and between trials. When fixation on the drift correction target was confirmed, the experimenter initiated the trial. Following Hermens and Walker (2015) experimental protocol and to minimize any memory effect resulted from the repetition of the same scenes, images were presented for 2000 ms. After 38 trials participants were offered a break and where needed a re-calibration was performed at the drift correction target. At the end of the experiment, participants were debriefed about the purpose of the study and were given the opportunity to ask questions about the experiment. Experimenter also asked participants few questions about the experiment (e.g., if they were aware of the cues, if cues where pointing at a certain location). None of these data were recorded and all these questions were part of participants' debriefing.

2.2.1.6 Data Analysis

Raw data from participants' eye movements were automatically parsed into fixations and saccade with the use of Eyelink's parser software (SR Research, Ontario, Canada). With the use of custom-built Matlab 2016a code, fixations were assigned to specific regions of interest (ROIs) by superimposing them onto colour coded versions of the images. These colour coded versions were created with the GNU image manipulation program (GIMP), with which each ROI was coloured with a specific colour. Based on Hermens and Walker

(2015)'s study, regions of interest were defined as actors' face, body and hand (for pointing cues) and arrow. In contrast to Hermens and Walker (2015), actors' entire body and not just the upper body was considered as a ROI. In addition, ROIs were not rectangular boxes surrounding the designated regions but tighter regions strictly defining each body part. For arrows cues ROI excluded the white background surrounding the arrow. Information about the ROIs' size per cueing condition can be found in Appendix (see Table A.1).

For most of the analyses linear mixed-effects model analyses were conducted with the lme4 package (Bates, Mächler, Bolker, & Walker, 2014) in R. Linear mixed-effects analyses have two important advantages: (1) the analysis can deal with missing data, which occurs when participants do not fixated a certain region of interest, (2) the analysis takes into account variability across participants (as in the more traditional ANOVAs and t-tests) and images (discounted in traditional ANOVAs and t-tests). For this model the statistics that is reported is the χ^2 values and p-value only (Baayen, Davidson, & Bates, 2008; Hermens, Golubickis, & Macrae, 2018; Magezi, 2015; Quené & Van den Bergh, 2008). Chi-square and p values express the statistical significance of the difference in goodness of fit between one model (e.g., one with an interaction and main effects) and a therein nested model (e.g., one with the main effect only). Examples of the model can be found in Appendix A.

To ensure that there was no effect on the results from the repetition of the same scenes two analyses will be presented: (1) one across all data for all participants, treating cues as repeated measures and (2) one analysing the first block of trials (first presentation of scenes) per participant, treating cues as a between-subjects factor. The latter analysis will be reported only for the dwell times on the cues and the saccades analyses (main analyses of this thesis). Bonferroni correction was applied where needed.

2.2.2 Results

2.2.2.1 Dwell Times on the cues

Cues for the gaze and pointing cues can be defined in different ways. This could involve the entire body of the actor, including the head and arm, or they could involve the main body part involved in the cueing (the arm for pointing and head for gazing). Based on Hermens and Walker (2015) analysis on the cues, the present data are analysed by including the entire body as a cue. Note that the dwell times on the cues are expressed as a percentage of the total trial time.

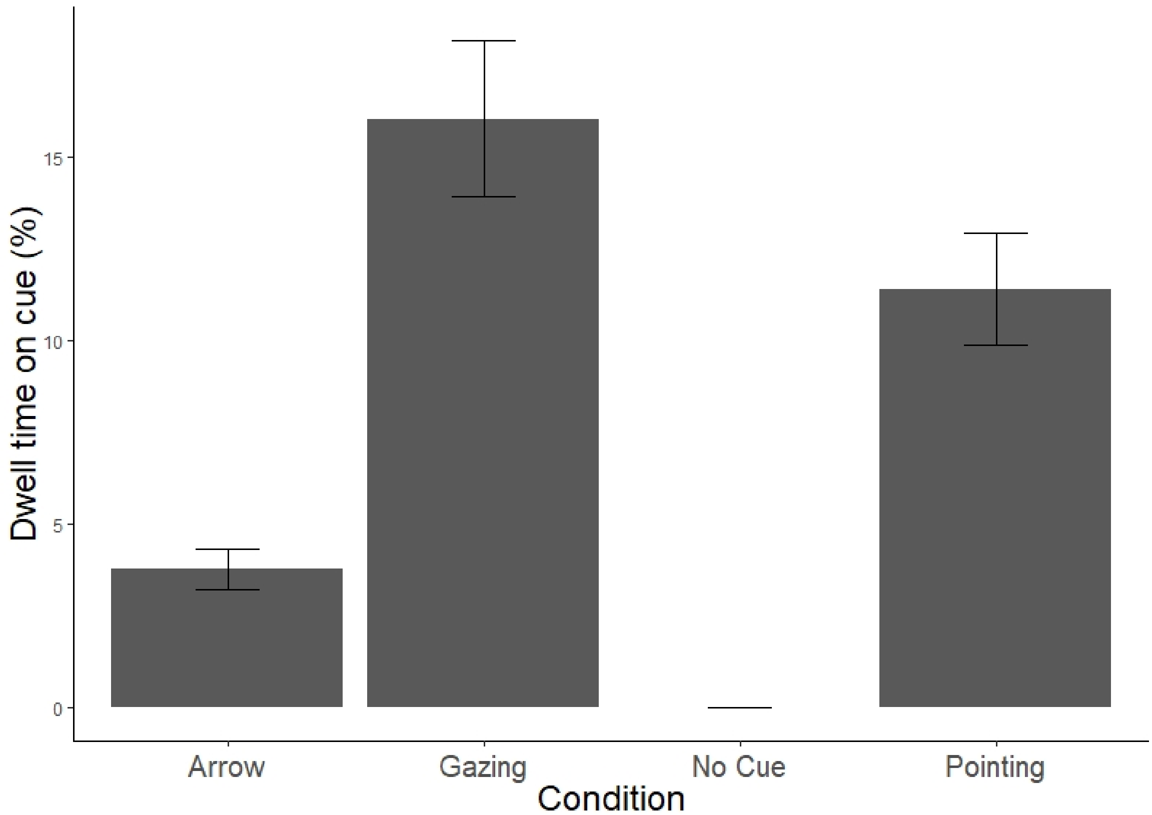

Fig. 2.2 Dwell Times on the cues for the different conditions. Averages were obtained by first pooling across scenes and then across participants. Error bars show the standard error across participants.

Figure 2.2 shows the dwell times on the cue for the different conditions, pooled across the scenes. Note that the 'No Cue' condition did not contain a cue and therefore has zero dwell times. The plot suggests that dwell times on gaze cues were longer than those on pointing cues and arrows. This observation is confirmed by a linear-mixed effects analysis, using participants and scenes as random factors, comparing a model with and without the condition parameter ($\chi^2(3) = 360.05, p < 0.0001$), indicating that cue dwell times differ significantly across cueing conditions. When only the first block of trials was analysed, similar results were found ($\chi^2(3) = 101.54, p < 0.0001$). This finding is not entirely surprising, as there is no cue in the 'No Cue' condition, and therefore dwell times for this condition are always zero. Therefore linear mixed-effects paired comparisons between the three cue-present conditions were conducted, showing longer cue dwell times for gazing than for pointing ($\chi^2(1) = 15.12, p < 0.001$), longer cue dwell times for pointing than for arrows ($\chi^2(1) = 48.85, p < 0.001$) and longer cue dwell times for gazing than for arrows ($\chi^2(1) = 104.99, p < 0.0001$). Similar results were found when the first block of trials was analysed (gazing vs pointing: $\chi^2(1) = 4.65, p = 0.03$;pointing vs arrow: $\chi^2(1) = 21.78, p < 0.0001$ and gazing vs arrow: $\chi^2(1) = 21.78, p < 0.0001$). These results contrast those by Hermens and Walker (2015) in that gazing cues attract more attention than pointing cues in the present study, while no difference was found between dwell times on gaze, pointing and arrow cues by Hermens and Walker (2015). A possible reason could be how regions of interest are defined. Hermens and Walker (2015) used large rectangular boxes around the arm, head and upper body for pointing cues and slightly smaller rectangular boxes around the head and upper body for gaze cues, whereas in the present study a tight region around the cues is used. This may also explain why arrow cues attracted less of the observers' gaze, because they were defined as just the arrow, not the entire sign.

2.2.2.2 Trials with fixations on the cue

Another measure of how strongly the cues attract attention is the percentage of trials with at least one fixation on the cue. If a cue attracts attention, a large percentage of trials with at least one fixation on that cue is expected. Figure 2.3. show that this percentage of trials with a fixation on the cue is lower for pointing (head or arm) than for gazing cues (linear mixed-effects logistic regression: $\chi^2(1) = 2.96, p = 0.003$) and lower for arrows than gazing cues ($\chi^2(1) = 15.1, p < 0.0001$). There are substantial numbers of trials without a fixation on the cue (around 30% for the gaze and pointing cues, and around 80% for the arrows).

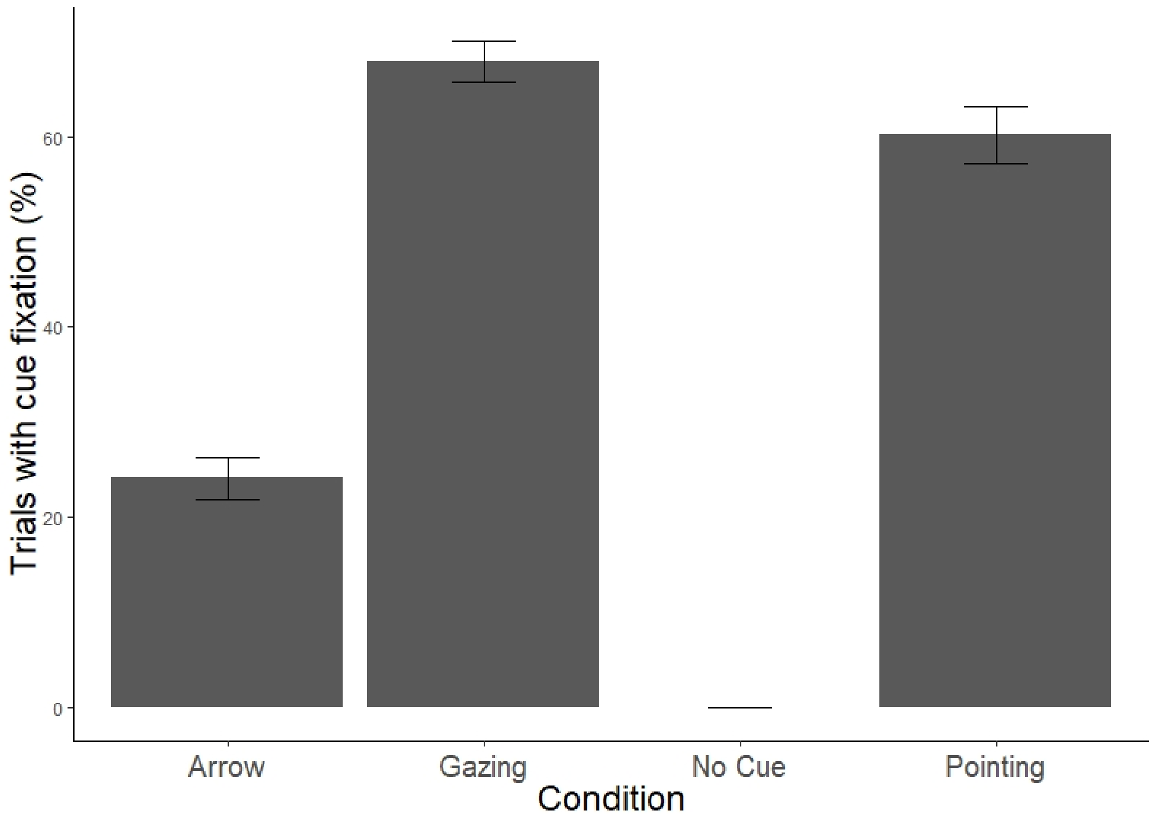

Fig. 2.3 Percentage of trials with fixations on the cues for the four conditions. Error bars represent the standard error of the mean across participants.

2.2.2.3 Dwell Times on Head, Body and Arm

In the analysis so far, the gaze and pointing cues were assumed to include the head and arm, as well as actors' entire body. To examine which region participants fixate for two types of cues, Figure 2.4 plots the dwell times for these separate regions. This plot shows the percentage of time (total) spent on each of the sub-regions, given a fixation on the region (person looked at the different body parts). This plot suggests that participants fixate the head most for both types of cues followed by the body and finally the arm for the pointing cues.

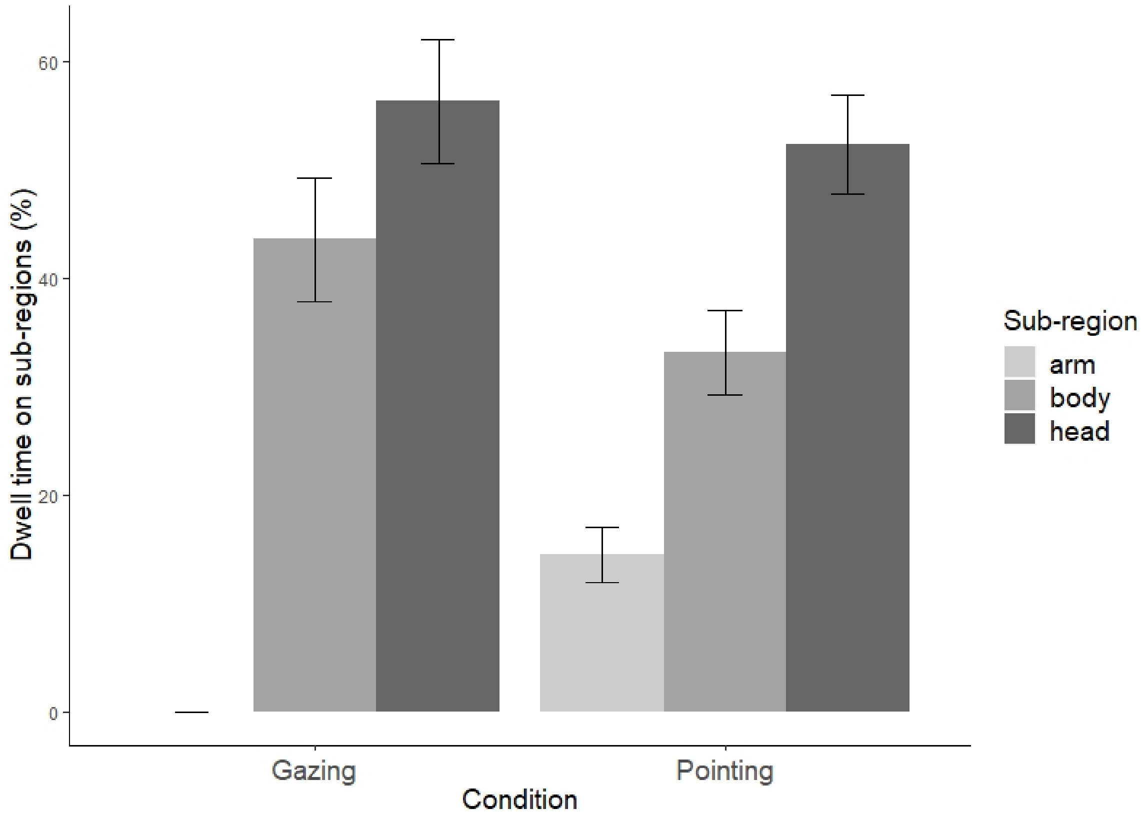

Fig. 2.4 Dwell time on the three sub-regions for the two social cues. Error bars represent the standard error of the mean across participants.

A linear mixed-effects model analysis (allowing for missing rows due to participants sometimes not fixating the people in the images) confirms these observations. The interaction analysis shows that the distribution of fixations across the three ROIs was different for gazing and pointing cues ($\chi^2(2) = 30.8, p < 0.0001$). However, this is not surprising as there is no arm in the gazing condition. When arm was removed from the analysis, significant interaction was found between the two body parts (body and head) and the two social cues ($\chi^2(1) = 0.94, p = 0.03$). Pairwise linear mixed-effects analyses examining the differences between the dwell times of the two regions showed shorter dwell times on the head for pointing cues ($\chi^2(1) = 5.33, p = 0.02$) and shorter dwell times on the body ($\chi^2(1) = 4.83, p = 0.03$).

Further set of pairwise comparisons on the differences between the body parts in each social cue was carried out. Linear mixed-effects model analysis on the three sub-regions in pointing shorter dwell times on the arm, compared to the other two sub-regions (arm vs body: $\chi^2(1) = 57.99, p < 0.0001$; arm vs head: $\chi^2(1) = 204.51, p < 0.0001$) and longer dwell times on the head when compared with body ($\chi^2(1) = 46.83, p < 0.0001$). Similar results were obtained for dwell times on the two regions for the gaze cues, which showed longer dwell times on the head than on the body ($\chi^2(1) = 48.94, p < 0.0001$).

2.2.2.4 Dwell Times on the target

Because of the possible confound between ROI size, ROI location and dwell times, dwell times on cues are difficult to interpret. In contrast, dwell times on cued objects are not expected to be influenced by such factors, as the cued object is always the same. Figure 2.5 shows the dwell times on the cued object ('target') for the four different conditions, pooled across the scenes and they are expressed as a percentage of the total trial time. Dwell times on the target suggest that in the presence of an arrows cue, targets were fixated for longer than in the presence of the social cues and the no-cue ('No Cue') condition.

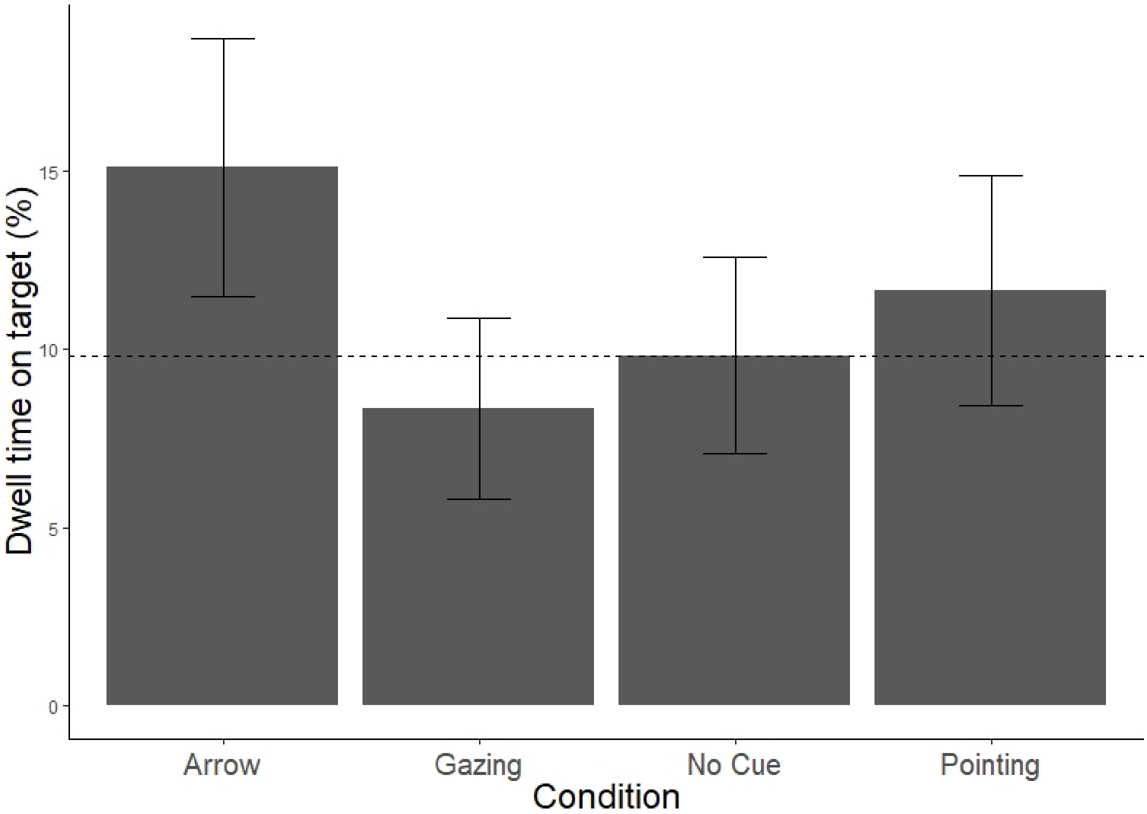

Fig. 2.5 Dwell Times on the target for the four conditions. Error bars represent the standard error of the mean across participants.

A linear mixed-effects analysis on the pooled dwell times on the targets shows a significant difference between the cueing conditions ($\chi^2(3) = 67.67, p < 0.0001$). Post-hoc analyses show significant differences in target dwell times between gazing and pointing ($\chi^2(1) = 15.03, p = 0.00011$), between arrows and pointing ($\chi^2(1) = 13.58, p = 0.0002$) and between arrows and gazing ($\chi^2(1) = 56.33, p < 0.0001$). These results suggest that targets were fixated longer for arrow cues than for the other conditions, that targets were fixated longer for pointing than for gaze cues. These findings may, in part, be explained by longer dwell times on the gaze cues, meaning there is less time available to fixate the cued object.

2.2.2.5 Trials with fixations on the target

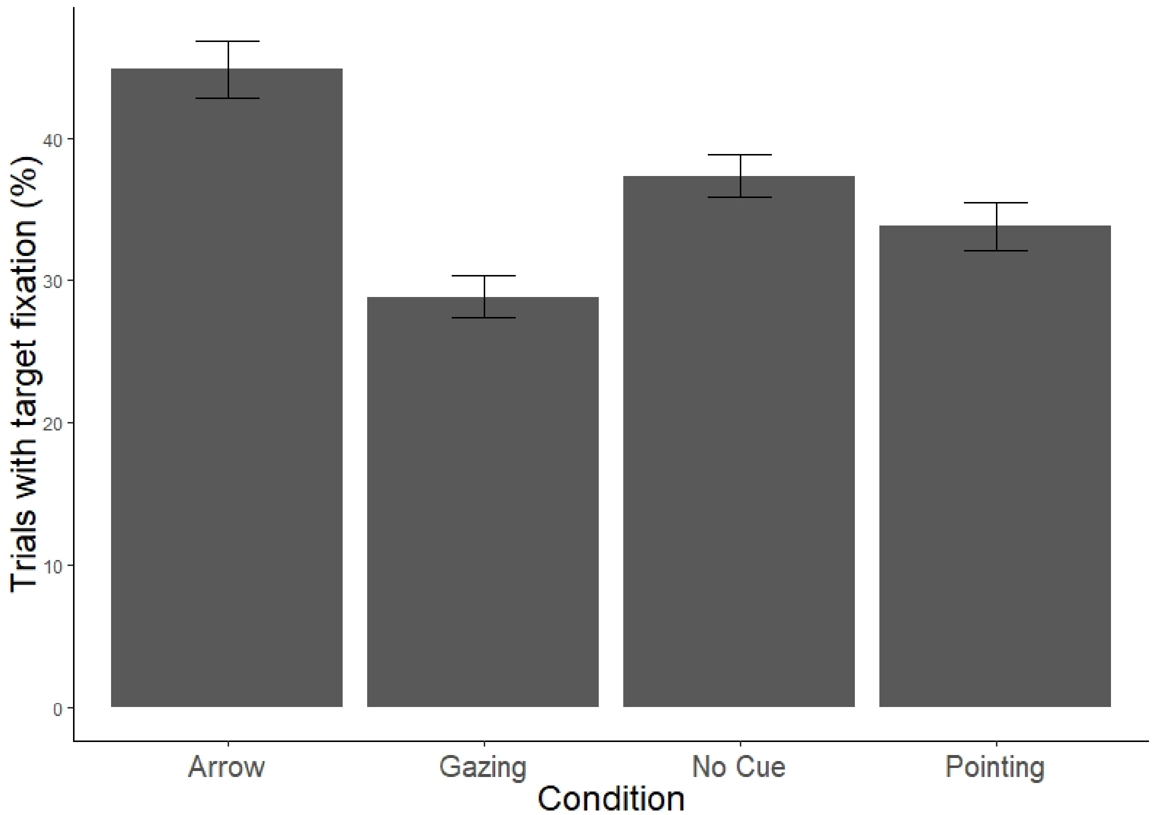

Fig. 2.6 Percentage of trials with target fixations for the four conditions. Error bars represent the standard error of the mean across participants.

The percentage of trials with at least one fixation on the target (Figure 2.6.) shows a similar pattern as the dwell times earlier, with the largest number of trials with a fixation on the target for arrows, followed by the no cue condition, and the gazing and pointing conditions. Linear mixed-effects logistic regressions shows significantly fewer trials with fixations on the target for gazing compared to no-cue trials ($\chi^2(2) = 4.06, p < 0.0001$), similar number of trials with fixations on the target for pointing and no-cue trials ($\chi^2(1) = 1.69, p = 0.091$) and significantly more trials with fixations on the target for arrows, compared to no-cue trials ($\chi^2(1) = 3.44, p = 0.001$).

2.2.2.6 Direction of saccades

The number of trials with fixations on the cue and the target is somewhat limited, but it still allows for an analysis of the direction of the saccades from the cue (i.e., where do the eyes go after fixating the cue), which provides the most direct measure of how strongly the cue shifts an observer's attention to the target (Hermens & Walker, 2015). Figure 2.7 shows that most saccades leave the cue, but do not go to the target, suggesting relatively weak cueing.

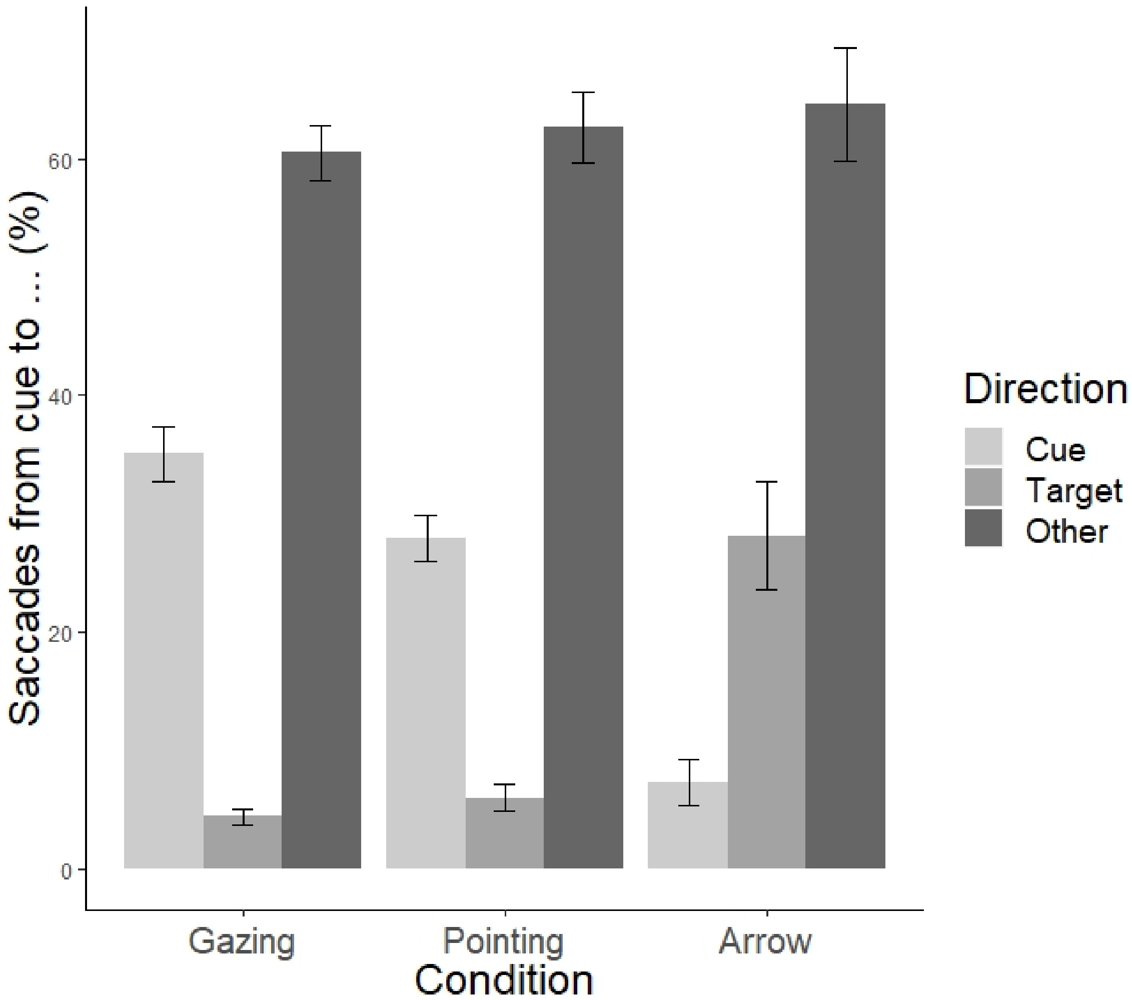

Fig. 2.7 Saccades leaving the cue to the three directions. Error bars represent the standard error of the mean across participants.

Whether just the first fixation on the cue is used, or all fixations on the cue, does not seem to make much of a difference (data not shown). In line with Hermens and Walker

(2015) all saccades are therefore used here. A linear mixed-effects analysis shows that the direction of the saccades depends on the cue (interaction between condition and cue direction, all saccades, all three directions ($\chi^2(4) = 83.53, p < 0.0001$). First block of trials analysis on the interaction between the condition and cue direction showed the same results ($\chi^2(4) = 15.41, p = 0.0004$). These differences reflect differences in saccades going back to the cue ($\chi^2(2) = 74.97 p < 0.0001$) and saccades going to the target ($\chi^2(2) = 38.00, p < 0.0001$), but not saccades going elsewhere ($\chi^2(2) = 0.76, p = 0.68$). Same results were obtained when only the first block of trials was used (saccades going back to the cue: $\chi^2(2) = 8.23, p = 0.02$; saccades going to the target: $\chi^2(2) = 5.01, p = 0.03$ and saccades going elsewhere: $\chi^2(2) = 6.33, p = 0.08$).

If only the saccades leaving the cue are considered, the direction also depends on the cue type ($\chi^2(2) = 33.76, p < 0.0001$), where target directed saccades ($\chi^2(2) = 20.68, p < 0.0001$) and elsewhere directed saccades ($\chi^2(2) = 14.15, p < 0.0001$) are affected by the type of cue. Same reuslts were found when only the first block of trials was analysed. Target directed saccades differ between arrows and pointing ($\chi^2(1) = 16.66, p < 0.0001$), between arrows and gazing ($\chi^2 = 19.08, p < 0.0001$), but not between pointing and gazing ($\chi^2(1) = 0.097, p = 0.76$). Same results were found when first block of trials was analysed (arrow vs pointing:$\chi^2(1) = 1.54, p = 0.02$; arrow vs gazing: $\chi^2(1) = 3.56, p = 0.04$ and gazing vs pointing: $\chi^2(1) = 2.90, p = 0.08$). The results match those by Hermens and Walker (2015) and show more cue following for arrows than for gazing. In contrast, arrows now also induce stronger cue following than pointing cues.

2.2.3 Discussion

Experiment 1 extended the work by Hermens and Walker (2015) and explored the effects of social and symbolic cues on observers' gaze behaviour when cues are presented in natural environments. Most work on social attention has presented cues embedded in a blank background and at the center of fixation (Frischen et al., 2007), and only recently studies (e.g., Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010), have moved towards presenting cues in the natural environment (in the form of photographs of natural scenes). The methodology used in the experiment was heavily influenced by Hermens and Walker (2015) and to an extent it tried to replicate the results of its free viewing task. Results show a similar pattern of results overall and thereby provide additional evidence on how social and symbolic cues differ in terms of attentional capturing and shifting.

Dwell times on the gaze cue were longer than those on the pointing and arrow cues, even if the body was included in the pointing cue. These results are largely in agreement with previous findings (e.g., Fletcher-Watson et al., 2008; Zwickel & Võ, 2010) but seem to deviate from those by Hermens and Walker (2015), where pointing cues received similar dwell times as gaze cues. This discrepancy in the results can be due to the different use of regions of interest (ROIs). In the present experiment ROIs were defined tightly to the actors' body, whereas Hermens and Walker (2015) used rectangular boxes that, particularly for the pointing cues, often containing part of the background. By defining ROIs tightly around the cues, a more precise analysis of eye movement was obtained. While Hermens and Walker (2015) suggested that the size of their regions of interest did not affect the results, recent work from Orquin, Ashby, and Clarke (2015) and the present study suggest that the size of a ROI can influence the outcome of a study.

Past studies have suggested that faces strongly attract attention (Birmingham et al., 2009), and a subsequent analysis was therefore performed to examine whether fixations on the gaze and pointing cues were indeed more often targeted to this region. The results showed that for both social cues, actors' face were fixated at most, with body and arm fixated less under the pointing condition. Although the head occupies a small area of the body, it is an important feature, which delivers a deeper and more complex message (e.g., people's intentions, feelings) as well as directing to a certain location (Bateson, Nettle, & Roberts, 2006; Csibra, 2010; Emery, 2000). These results for the head are not surprising, as they agree with previous literature (e.g., Birmingham et al., 2009; Hermens & Walker, 2015) that also presented full body actors instead of disembodied heads. Pointing hands seems to be overshadowed by gazing heads, when presented simultaneously.

Dwell times on the cue and on the cued object may not be sufficiently informative about how strongly a cue shifts attention to the cued object. The present results show that dwell times on the cue were longest for the gaze cue, but that dwell times on the cued object were shortest. A reason is that participants were busy fixating the cue and therefore did not have time left to fixate the cued object. For this reason, saccades following a fixation on the cue were analysed for their direction (see also Hermens & Walker, 2015). This analysis showed that arrows led to stronger cueing than the two social cues, in agreement with some past findings (Brignani, Guzzon, Marzi, & Miniussi, 2009; Kuhn & Kingstone, 2009). However, arrows on these studies did not compete with social cues produced by actual people, positioned in images of natural scenes. More recent studies (e.g., Hermens & Walker, 2015) have suggested that pointing hand is the one which induce a better cueing effect. Results for arrows cueing effect in the present study, comes in contrast with part of the previous literature (e.g., Burton et al., 2009; Hermens & Walker, 2015; Langton & Bruce, 2000). In these studies when peripheral social or symbolic cues were presented, pointing cues demonstrated a stronger cueing effect.

But what made arrows in our case to be more successful? One difference between the arrow cue in the present and the study by Hermens and Walker (2015) is the size of the arrow. Because the arrows occupied a much smaller region of the image in the Hermens and Walker (2015) study than the gaze and pointing cues (particularly if the body region is taken into account), the size of the arrows was increased from A4 to A3 in the present study. This small change might have allowed the arrows to be more prominent and given them the opportunity to equally compete social cues. In addition, by making arrows more salient from the one

Hermens and Walker (2015) study used, by defining more precise regions of interest and by presenting them not directly at fixation (similar to Hermens & Walker, 2015) might have triggered a more effective orientation effect from the two social cues.

When in the presence of a social cue, observers might find it difficult to disengage and follow their direction when compared to the arrows. This is reflected in the analysis of the saccades and dwell times, where social cues collected the majority of fixations and observers spend a lot of time revisiting the cues. One can suggest that social cues did not produce a better cueing effect due to the fact that the observers spent more time distinguishing the actor's facial expression, characteristics or intentions. Resulting to a delayed disengagement and attentional shift to the targeted object. The study by Senju and Hasegawa (2005) can support this claim, as they suggested that face stimuli with direct direction towards the target, can delay disengagement of attention to a target. In contrast to social cues, arrows may be easier to interpret, as the message conveyed by them is clear and concise (pointing to a direction).

While participants often looked at the cues and cued objects, they also spent a substantial portion of each trial looking at the background. Compared to past studies (e.g., Hermens & Walker, 2015) the present study showed relatively long dwell times on the background. A possible reason could be the relatively dense and complex backgrounds used in the present experiment. As cues were part of a free viewing task and no instruction were provided to bias observer towards them, other elements of the scene drew the observers' attention away from the cues. Another potential explanation is that when cues are presented in natural scenes the direction they point at is ambiguous. Past studies have suggested that participants have difficulties estimating the direction of a cue in a photograph of a 3D scene (Doumen, Kappers, & Koenderink, 2010). Participants may therefore have understood the direction indicated by the cues, but have looked at non cued items (background), because this was what they thought the cue was pointing or looking at. Great care was taken to take photographs with little confusion about where the cues were pointing, but not all photographs may have succeeded. A follow-up questionnaire could therefore ask participants to indicate (e.g., with a mouse-click) which object each cue was pointing at to verify whether the intended target was indeed perceived as such.

Results from dwell times and saccades are to an extent in agreement with the previous studies (Fletcher-Watson et al., 2008; Hermens & Walker, 2015; Zwickel & Võ, 2010) where a single informative to targeted object location cue presented in the scene. As a result, this might have made cues more notable and one might ask if similar effects will occur when multiple social and symbolic cues are presented in the scene and if the direction of these cues (e.g., in congruent or competition) will influence these effects. Therefore, Experiment 2 in this thesis aims to address this question by presenting social and symbolic cues either in competition (e.g., one cue is pointing at the targeted item and another one at an opposite direction) or congruent (e.g., both cues are pointing at the same targeted item).

2.3 Experiment 2

Experiment 2 examined where people look when multiple social and symbolic cues are presented in an image. Cues are placed either in direct competition (pointing at different targets) or in congruent (pointing at the same target).

2.3.1 Methods

2.3.1.1 Participants

For Experiment 2, twenty-five students from the University of Lincoln were recruited (Males = 9, aged between 19 and 30). All participants provided a written consent form approved by the local ethics committee and all reported normal or corrected to normal vision. As a reimbursement for their participation each received course credit. The study was approved by the from University of Lincoln (Psychology department) Ethics Committee.

2.3.1.2 Apparatus

Apparatus for Experiment 2 was the same as for Experiment 1.

2.3.1.3 Stimuli

Fig. 2.8 Example of congruent and in competition stimuli used in Experiment 2. The left image shows (top) one actor pointing at the target's location with other gazing at the same location. The right image (top) shows one actor gazing at the target's location with second actor pointing away from the target. Left image (bottom) shows one person gazing at the target's location with an arrow pointing at the same direction. Right image (bottom) shows one actor pointing at the direction of the target with a second actor gazing away from the target.

Seventy photographic images of seven natural scenes were used. Images showed indoors and outdoors areas of the University of Lincoln campus. Similarly to Experiment 1 and for the same reasons, images were scaled down to the size of 800 x 600 pixels (approximately 16°

x 12° of visual angle) using the IrfanView software. In each image different combinations of congruent (3 combinations) and in competition (6 combinations) cues were presented (Table 2.1). In addition, each scene was photographed without any cues. Cues were presented bilaterally of targeted object (Figure 2.8). Cues' location varied to avoid participants familiarize and predict where these cues might occur, and thereby avoid anticipatory scan paths. Note that this study is the first to vary the cue locations. None of the previous studies (e.g., Birmingham et al., 2009) using multiple cues, have manipulated cues' location.

Table 2.1 Combination of congruent and in competition cues presented in Experiment 2

Combination	Cue 1	Cue 2
Congruent	Arrow	Point
	Arrow	Gaze
	Gaze	Point
Competition	Arrow Towards	Point Away
	Arrow Towards	Gaze Away
	Arrow Away	Point Towards
	Arrow Away	Gaze Towards
	Point Towards	Gaze Away
	Point Away	Gaze Towards

2.3.1.4 Design

Design for Experiment 2 was the same as for Experiment 1.

2.3.1.5 Procedure

After receiving instructions and signing the consent form, participants placed their head in the chin rest and the 9-point calibration of the Eyelink 1000 system was performed. Participants then performed the 70 trials of the experiment, each of which started with a drift correction.

Similar to Experiment 1 and to Hermens and Walker (2015) study, images were presented for 2000 ms. Drift correction was placed to the periphery of the image to avoid any effect from the central bias. Participants were given the opportunity to take a break after 35 trials. After completing the trials, participants were debriefed and were given the opportunity to ask questions.

2.3.1.6 Data Analysis

As in Experiment 1, same regions of interested were created with GIMP software and statistical analyses were conducted using R (version 3.4.1, R Development Core Team 2017). For the statistical analysis of the data, linear mixed-effects models instead of a more traditional analyses (repeated measures ANOVA, paired sample t-tests) was used. Similar to Experiment 1, two analyses will be reported (analysis of the overall data and first block of trials analysis). First block of trials analysis will be reported only for the two key analyses of this thesis (dwell times on the cues and saccades analyses) to show that repetition of the images did not affect the results. When multiple comparisons were performed, the significance level (critical p-value) was corrected with a Bonferroni correction.

2.3.2 Results

2.3.2.1 Dwell Times on congruent cues

The analysis of the data was similar to that employed in Experiment 1. When computing the dwell times on cues, the actors' bodies were considered part of the cues. Figure 2.9 shows the dwell times on the cues when they were both pointing at the target, with the averages computed across the scenes in which the cues appeared. Note that the dwell times presented in this graph and the following analyses, want to explore when people are looking at one of the two cues, which cue do they look at the most.

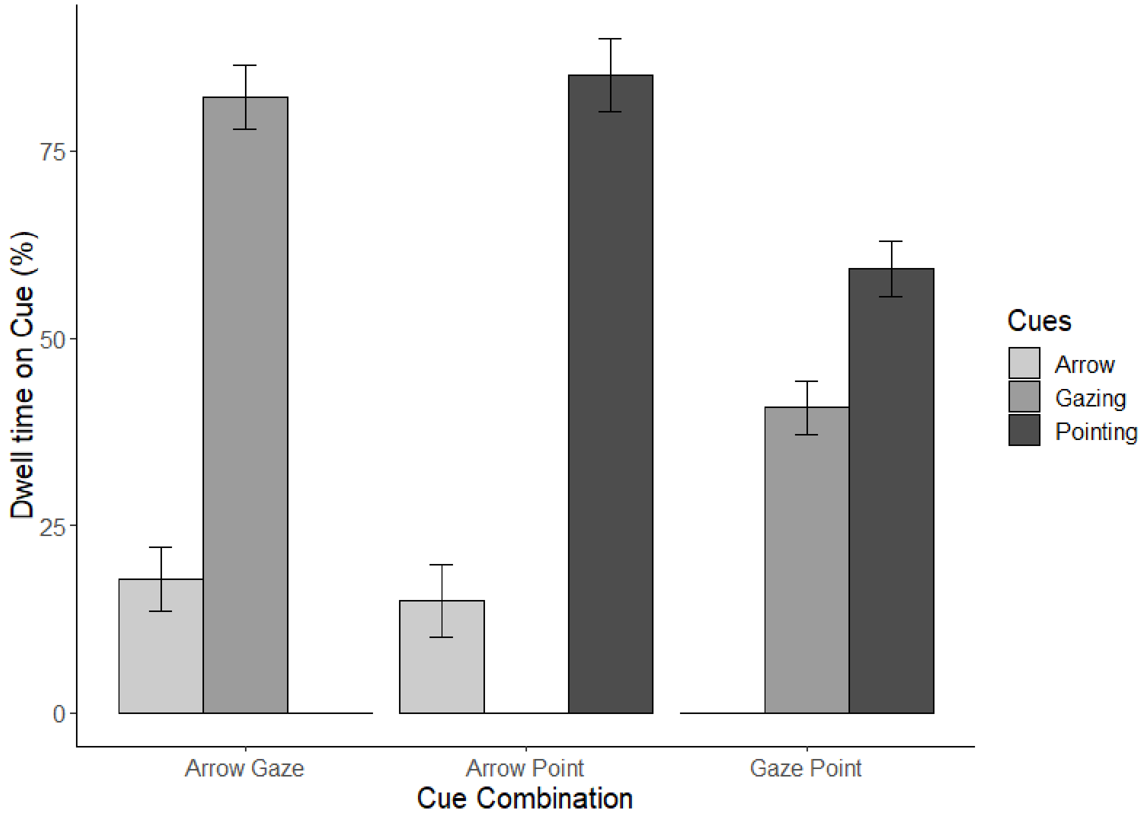

Fig. 2.9 Dwell Times on cues when both cues pointed towards the target. Gazing = an actor gazing at the target. Pointing = an actor pointing at the target. Error bars represent the standard error of the mean across participants.

Figure 2.9 suggests that social cues were looked the most when accompanied by another symbolic cue of the same direction. This graph also suggests that when pointing and gazing pointed at the same direction pointing received the majority of fixations compared to gazing. These results were confirmed by linear mixed-effects model analysis where gazing and pointing cues showed a significant difference in dwell times when combined with an arrow (Arrow Point: $\chi^2(1) = 433.52, p < 0.0001$; Arrow Gaze: $\chi^2(1) = 301.9, p < 0.0001$). The social cues were looked at most and the pointing cue was looked more when combined with a gazing cue (Gaze Point: $\chi^2(1) = 147.92, p < 0.0001$). When first block of trials was analysed similar results were obtained (Arrow Gaze: $\chi^2(1) = 30.59, p < 0.0001$; Arrow

Point: $\chi^2(1) = 63.92, p < 0.0001$ and Gaze Point: $\chi^2(1) = 24.95, p < 0.0001$) It should be noted that these results include actors' body and head (for gazing and pointing cues).

2.3.2.2 Trials with fixations on the congruent cues

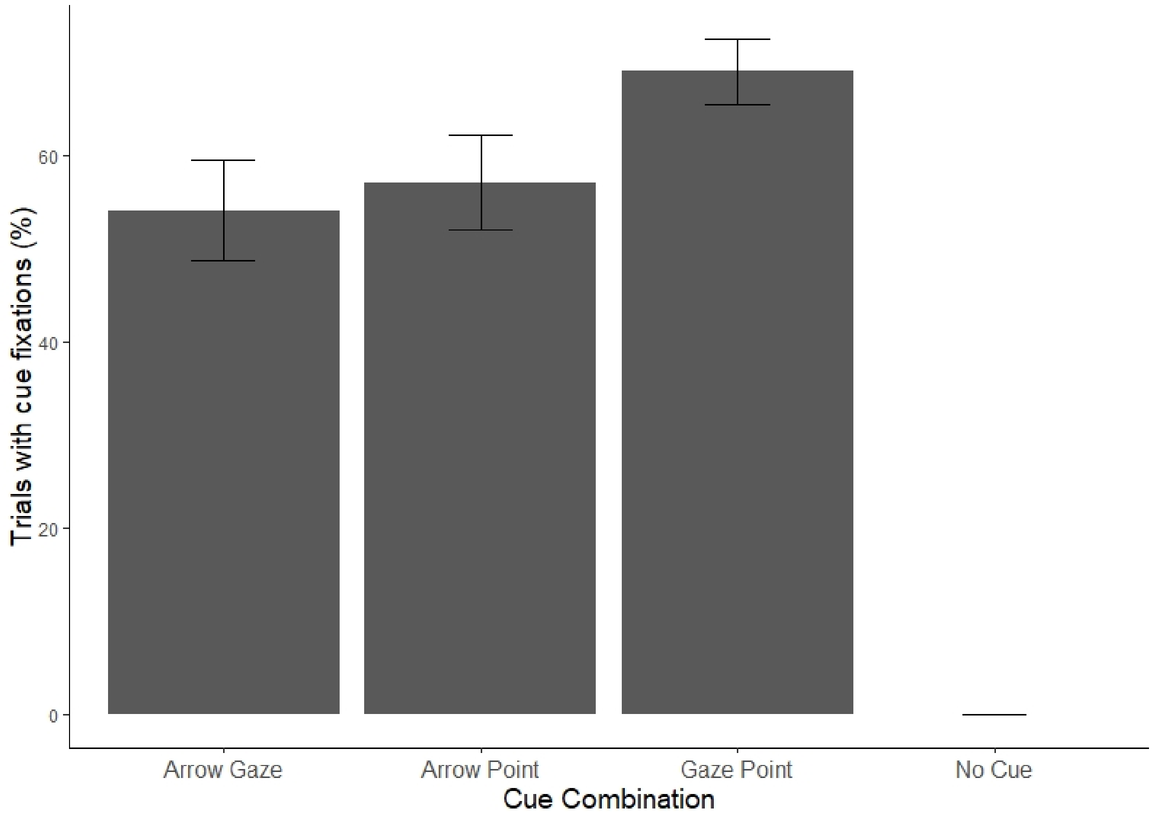

Fig. 2.10 Percentage of trials with fixations on the cues. Error bars represent the standard error of the mean across participants.

Similar to Experiment 1 another measure to explore how strongly cues attract attention is the percentage of trials with fixations on the cue combinations. Figure 2.10 shows that the percentage of the trials with fixations on the cue combination were lower for the combinations which contained a symbolic and social cue, higher for the combination with two social cues (Linear mixed-effects logistic regression: Arrow Point vs Gaze Point: $\chi^2(1) = 5.58, p = 0.02$; Arrow Gaze vs Gaze Point: $\chi^2(1) = 8.62, p = 0.003$) and difference between the combination with mixed cues (Linear mixed-effects logistic regression: Arrow Gaze vs Arrow Point: $\chi^2(1) = 0.36, p = 0.55$).

2.3.2.3 Dwell Times on Head, Body and Arm for congruent cues

When computing dwell times on social cues, the actors' bodies were included for gaze and pointing cues. Most studies suggest that when presented with natural scenes with actors, observers mostly look at their heads. However, the body and arm may here provide direction cues and may therefore also be looked at. To examine whether this was indeed the case, Figure 2.11 shows the dwell times on the three sub-regions (head, body, and arm) for the gazing and pointing cues in each combination consisting of social cues both pointing towards the target. The dwell times on the three body parts for gazing and pointing are calculated on how much time (total) spend on the person (given that they looked at the different body parts).

The Graph suggests larger dwell times for the head region only when the social cue was presented alongside with an arrow. A linear mixed-effects analysis confirmed an interaction between the two conditions (gazing, pointing) and the three sub-regions ($\chi^2(4) = 44.04, p < 0.0001$). This is not surprising as there was no hand in the gazing cue. When arm was removed, linear mixed-effects analysis on the interaction between the two social cues and the two body parts showed significant results ($\chi^2(3) = 5.97, p = 0.04$). Linear mixed-effects analysis on the main effect of the three sub-regions showed a significant difference for each region in each cue combination ($\chi^2(2) = 76.9, p < 0.0001$). Similar significant results were found when arm was removed from the analysis ($\chi^2(1) = 9.04, p = 0.003$). Detailed analysis was carried out for the three sub-regions in each cue combination showing (Arrow Gaze: $\chi^2(2) = 89.29, p < 0.001$; Arrow Point: $\chi^2(2) = 11.98, p = 0.002$; Gaze Point: $\chi^2(2) = 0.10, p = 0.77$). Subsequent linear mixed-effects analysis in each combination of

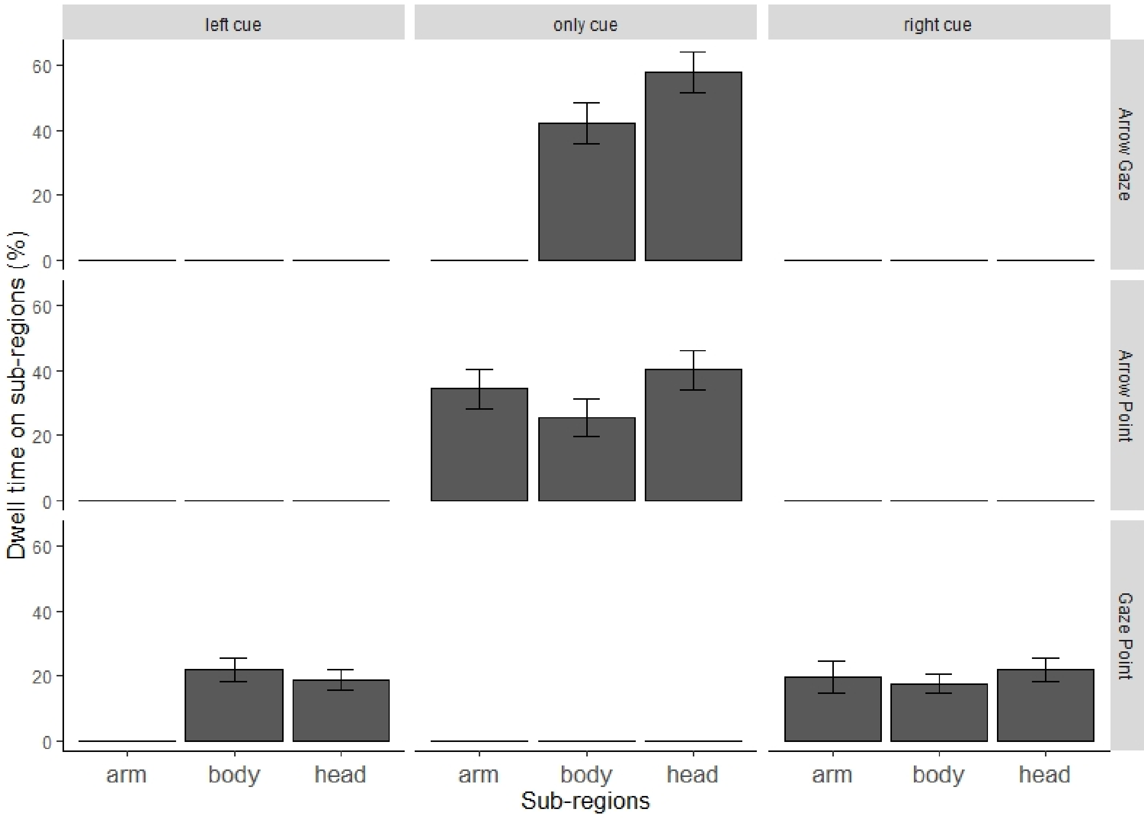

Fig. 2.11 Dwell times on the three body parts for the two social cues. The left part of the graph shows dwell times on the body sub-regions when social cue (gazing) was in combination with a pointing cue. The middle part of the graph shows dwell times on the sub-regions when social cue was in combination with an arrow. The right part of the graph shows dwell times on the sub-regions when one of the social cue (pointing) was in combination with gazing. Error bars represent the standard error of the mean across participants.

cues showed that heads collected the majority of fixations for the cue combinations containing a social cue (Table 2.2).

Table 2.2 Statistics for comparisons of dwell times on the three sub-regions across the three cue combinations of Experiment 2

Cue	Combination	χ^2-value	p-value
Arrow Point	arm vs body	0.85	0.42
	head vs body	7.79	0.01
	arm vs head	3.56	0.05
Arrow Gaze	head vs body	3.14	0.04
Gaze Point	head vs body	0.10	0.77

2.3.2.4 Dwell Times on the target for congruent cues

Figure 2.12 shows the dwell times on the target when there were two congruent cues present, suggesting that under the arrow with gaze condition the target was looked less than for other cue combinations. Interestingly, targets under the no-cue condition were looked at most. Moreover, arrows with pointing and gaze with pointing combinations lead to similar dwell times on the target. Dwell times on the targets are expressed as a percentage of the total trial time.

These observations from Figure 2.12 were confirmed by linear mixed-effects model analysis, showing significant differences in dwell times on the target across the cue combinations ($\chi^2(3) = 29.09, p < 0.0001$). As with the gaze behaviour on cues this is not surprising as targets were cued by two informative cues. Pairwise comparison analyses showed a significant difference between the arrow with gaze combination and two other conditions (vs arrow point: $\chi^2(1) = 4.38, p = 0.04$; vs no cue: $\chi^2(1) = 25.14, p < 0.0001$) and no significant difference compared to the gaze point condition ($\chi^2(1) = 3.63, p = 0.06$). Comparison between arrow with pointing and gaze with pointing cues showed no significant difference in dwell times on the target ($\chi^2(1) = 0.11, p = 0.74$). Compared to the no-cue condition, these

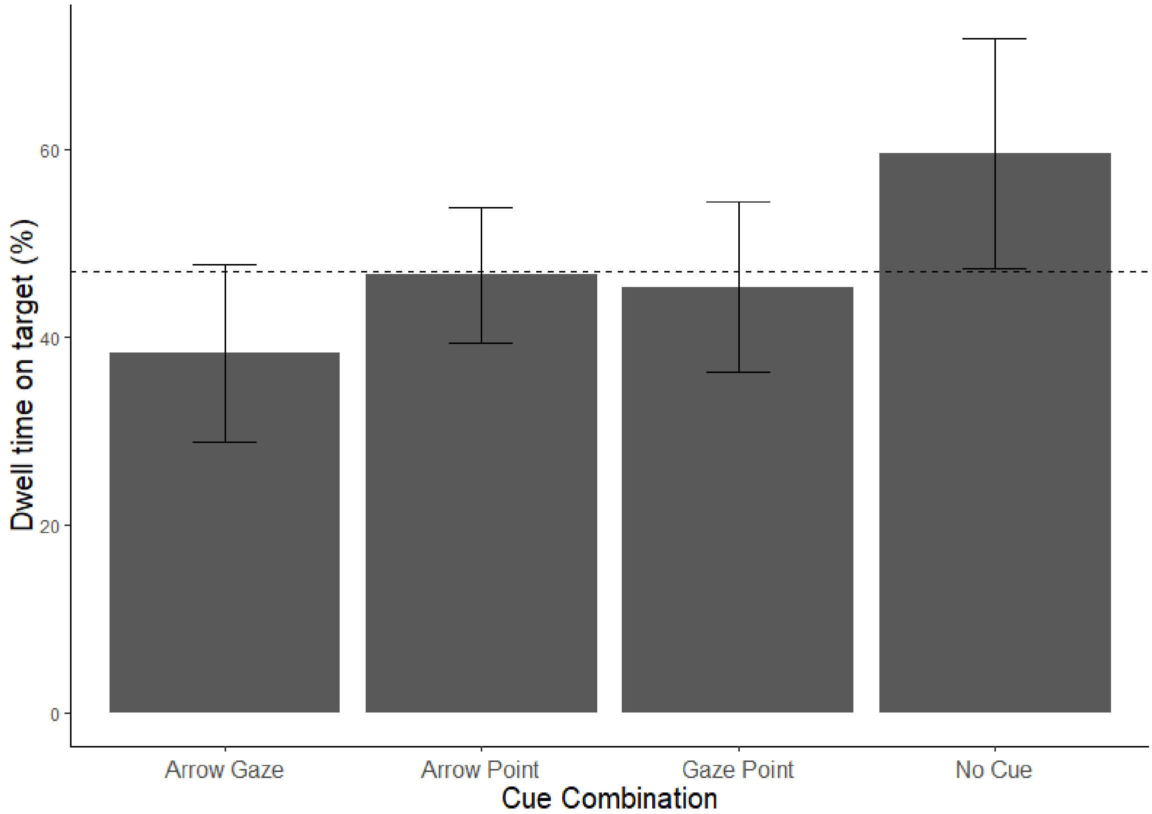

Fig. 2.12 Dwell times on target for cues under congruent condition.Error bars represent the standard error of the mean across participants.

combinations led to significantly shorter dwell times on the target (compared with gaze point: $\chi^2(1) = 12.64, p = 0.0003$ and compared with arrow point: $(\chi^2(1) = 9.54, p = 0.002)$.

2.3.2.5 Trials with fixations on the the target for congruent cues

Similar to Experiment 1 another measure to explore how targets attract participants' attention is the percentage of trials with fixations on the cue combinations. The percentage of trials with a fixation on the target (Figure 2.13) shows a similar pattern to the dwell times on the target shown earlier, with no difference on the percentage of the trials on the target for any of the cue combinations. Linear mixed-effects logistic regression confirmed this, showing no significant difference between the trials with fixations on the target $(\chi^2(3) = 7.08, p = 0.07)$.

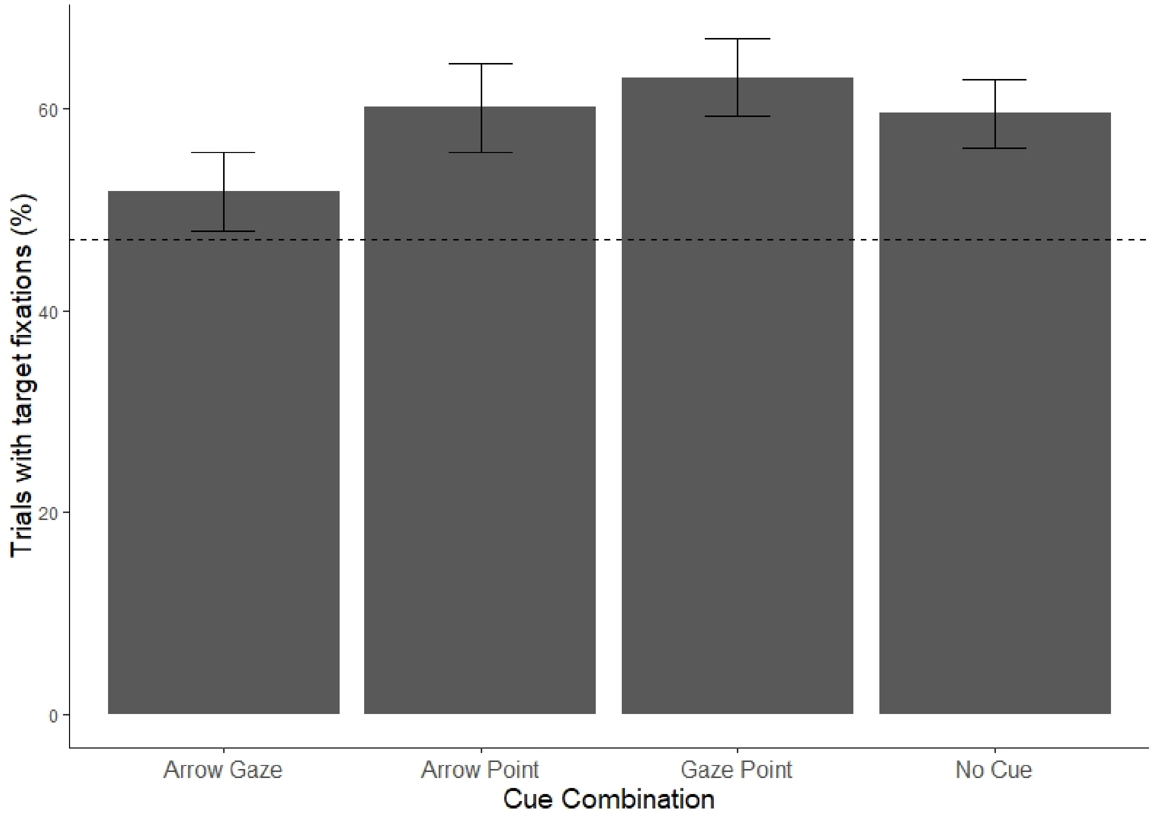

Fig. 2.13 Percentage of trials with fixations on the target per cue combination. Error bars represent the standard error of the mean across participants.

2.3.2.6 Direction of saccades for congruent cues

There were substantial numbers of trials with fixations on the cues and target. This allows for an exploration of the direction of saccades leaving each type of cue. Figure 2.14 suggests that arrows led to more saccades to the target when it was combined with a gaze cue.

A linear mixed-effects analysis shows that the direction of the saccades depends on the cue combination (interaction between cue combination and direction, all saccades, all three directions $\chi^2(3) = 2710, p < 0.0001$). When only the first block of trials was analysed similar reuslts were found ($\chi^2(3) = 458.22, p = 0.001$). If only the saccades leaving the cue are considered the target directed saccades are significantly different for the arrows, when paired with a gazing cue also pointing to the target ($\chi^2(1) = 5.69, p = 0.02$). The target directed saccades are not significantly different for pointing cue when paired with another social cue

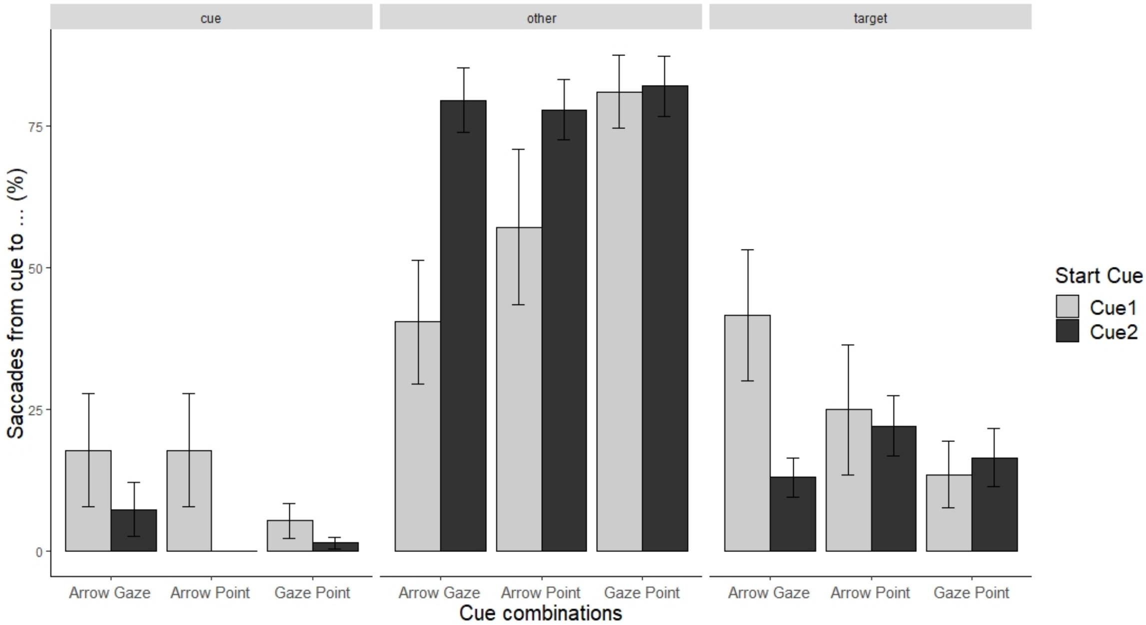

Fig. 2.14 Saccades to the other cue (left), elsewhere (middle) and the target (right) when two cues were both pointing to the target. Error bars represent the standard error of the mean across participants.

and an arrow pointing at the same direction (arrow vs point: $\chi^2(1) = 0.014, p = 0.90$; gaze with point: $\chi^2(1) = 0.21, p = 0.64$). Similar results were obtained when first block of trials was analysed (arrow vs point: $\chi^2(1) = 0.58, p = 0.65$; arrow vs gaze: $\chi^2(1) = 1.45, p = 0.03$ and gaze vs point: $\chi^2(1) = 1.18, p = 0.89$). Analysis of the difference in saccades going to 'elsewhere' (i.e., not to the other cue or the target) showed no significant difference between cue combinations ($\chi^2(2) = 0.40, p = 0.82$). Similarly, analysis of the number of times participants saccades to the same or the other cue showed no significant difference between the cue combinations ($\chi^2(2) = 2.76, p = 0.25$). Similar results were found when only the first block of trials was analysed.

2.3.2.7 Dwell Times on cues in competition

Figure 2.15 shows dwell times on cues under the contrasting condition, where one of the cues was pointing away from the target. Specifically, Figure 2.15 shows how attention is distributed across the two in competition cues in the scene. From the graph we see a pattern suggesting that social cues tend to capture observers' attention for longer than the arrows. As this pattern is the same for non-competing cues, it therefore suggests that the advantage for social cues is independent of the social cue's direction. Note that dwell times presented in this graph want to explore when people are looking at one of the two cues, which cue do they look at the most.

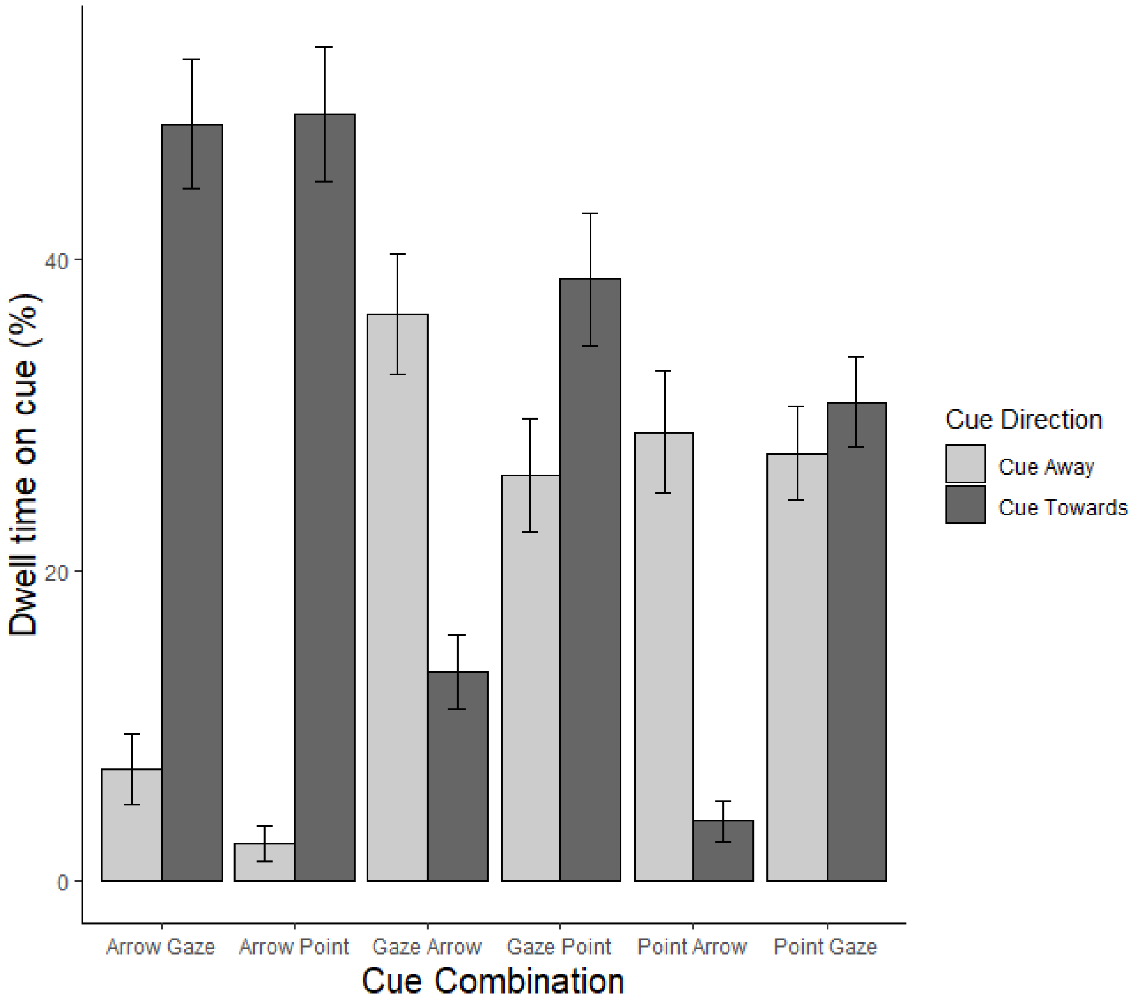

Fig. 2.15 Dwell times on competition cues (i.e., when the secondly listed cue was pointing away from the target, and the first of the listed cues was pointing towards). Generally, dwell times were longer on the social (gaze and pointing) cues than on the arrows, independent of whether they were pointing at the target. Error bars represent the standard error of the mean across participants.

This suggested advantage of social cues was explored with a linear mixed-effects model analysis (using participants and scenes as random factors) computing the interaction between the direction of each cue and the cue of a different direction. Interaction analysis on the arrow in relation to its direction and the second cue showed no significant difference ($\chi^2(1) = 0.03, p = 0.85$). Similar results were found when only the first block of trials was analysed ($\chi^2(1) = 0.50, p = 0.48$). Exploring the main effect for the arrows' direction showed a significant difference ($\chi^2(1) = 55.79, p < 0.0001$) and a significant main effect for the second cue ($\chi^2(1) = 10.67, p = 0.001$), suggesting that arrows direction towards and away the target and the cues paired with did not influence people's attention towards them. Similar results were found for the main effect of the second cue ($\chi^2(1) = 3.09, p = 0.03$) when first block of trials was analysed but not for the main effect for the arrow's direction ($\chi^2(1) = 1.32, p = 0.25$).

A linear mixed-effects models analysis showed a significant interaction between the direction of the pointing cue and the type of the second cue (arrow or gazing) ($\chi^2(1) = 29.77, p < 0.0001$). When only the first block of trials was analysed similar results were found ($\chi^2(1) = 4.52, p = 0.03$). No main effect was found for the second cue ($\chi^2(1) = 3.83, p = 0.06$) and a significant main effect of the direction of the cue ($\chi^2(1) = 83.76, p < 0.0001$). Similar results for the main effect of the direction of the pointing cue ($\chi^2(1) = 8.81, p = 0.003$) and the main effect of the second cue ($\chi^2(1) = 0.34, p = 0.56$) were found when the first block of trials was analysed. Finally, a linear mixed-effects models analysis showed a significant interaction between the direction of the gazing cue and the type of second cue ($\chi^2(1) = 43.30, p < 0.0001$). When only the first block of trials was analysed similar results were found ($\chi^2(1) = 18.20, p < 0.0001$). Significant results were found for the main effect for gazing cues direction ($\chi^2(1) = 48.98, p < 0.0001$) but no significant results for the second cue ($\chi^2(1) = 0.13, p = 0.71$). Similar results for the main effect of the direction of the gazing cue ($\chi^2(1) = 4.92, p = 0.03$) and the main effect of the second cue ($\chi^2(1) = 1.31, p = 0.25$)

were found when only the first block of trials was analysed. Table (2.3) shows the results from the linear mixed-effects analysis when exploring the differences on cues' dwell times for each cue combination. Analyses suggest that arrows failed to capture observers' attention when combined with another social cue. On the other hand, social cues were equally looked but only for an informative to the target direction gaze, whereas pointing was looked more than the gazing when it was informative to targets direction. When the first block of trials was analysed, similar results were found.

Table 2.3 Statistics for comparisons of dwell times on the cues across the different combinations of Experiment 2

Target Cue (towards)	Second Cue (away)	χ^2-value	p-value
Gaze	Arrow	87.82	0.0001
Point	Arrow	116.38	0.0001
Gaze	Point	0.52	0.47
Point	Gaze	7.16	0.007
Arrow	Gaze	28.38	0.0001
Arrow	Point	44.08	0.0001

2.3.2.8 Trials with fixations on the cue in competition

Figure 2.16 shows the cues were quite often not looked at. It plots the percentage of trials with at least one fixation on either cue. It also shows that cues were looked at more for combinations that contained a social cue pointing or looking at the target and less when the arrow was pointing at the target. A linear mixed-effects analysis comparing the percentage of trials with fixations across cue combinations showed a significant effect of cue combination ($\chi^2(5) = 76.62, p < 0.0001$). The pattern matches the one of the dwell times on the cues and trials with informative arrows, showing less percentages of trials with at least one fixation to the cue (for detailed comparisons see Apendix A).

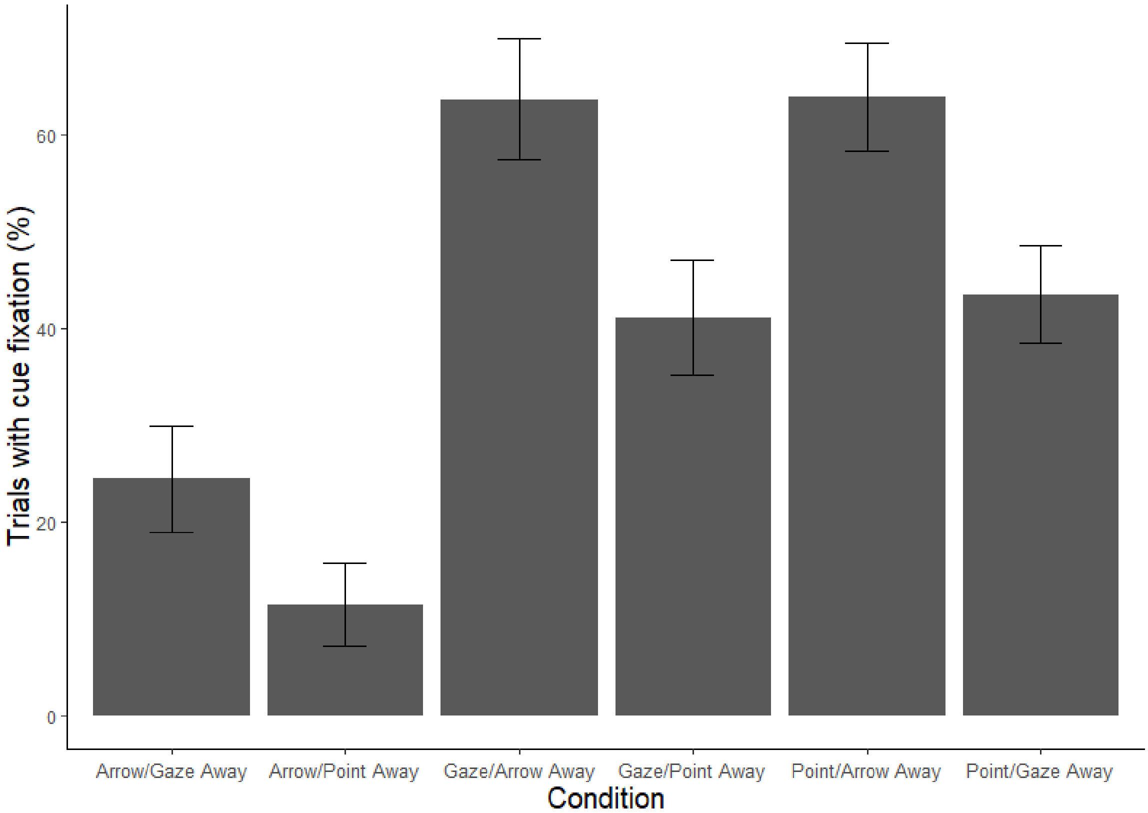

Fig. 2.16 Percentage of trials with fixations on the cues. Error bars represent the standard error of the mean across participants.

2.3.2.9 Dwell Times on Head, Body and Arm for cues in competition

In the analysis so far, gaze and pointing cues contained the body of the actor. To examine how strongly the body contributed to dwell times on the cues, Figure 2.17 shows the dwell times on the three body parts (head, arm and body) for each cue combination and direction. The dwell times on the three body parts are calculated on how much time (total) spend on the person (given that they looked at the different body parts). Graph suggest large dwell times on the head but only for the pointing cues pointing to the direction of the target.

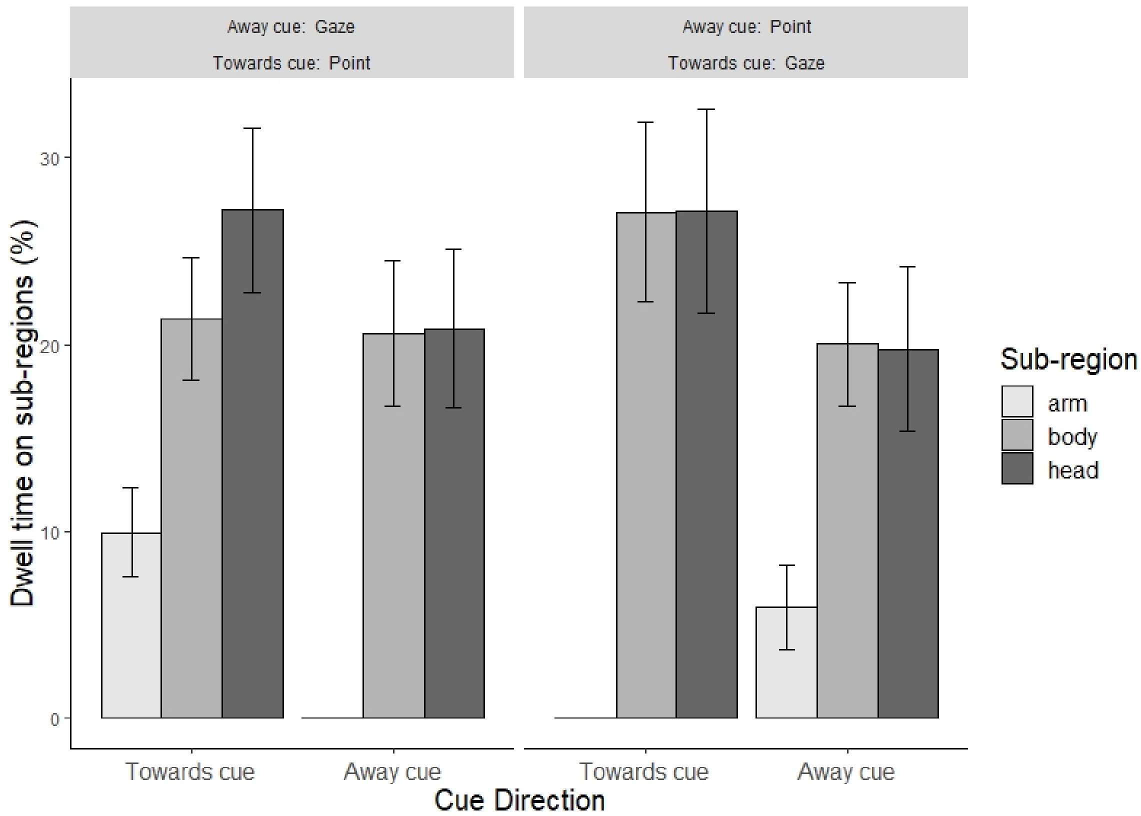

Fig. 2.17 Dwell Times on the three sub-regions (arms, body and heads) for each social cue, in each configuration with another cue. Error bars represent the standard error of the mean across participants.

The interaction between dwell times on the two body parts (body and head) and the direction of the gazing and pointing cues was not significant ($\chi^2(2) = 1.54, p = 0.21$), suggesting that the direction of the cues (away or towards the target) did not influence the distribution of dwell times on these three body locations. There was no main effect of the two body part both for the towards and the away directions (towards direction: $\chi^2(1) = 0.76, p = 0.38$; away: $\chi^2(1) = 0.84, p = 0.36$). Similar findings were found for the main effects of the body parts for each cue and for each direction (Table 2.4).

Table 2.4 Statistics for the main effect of the two body parts on the gazing cues and the three body parts on the pointing cues across the different directions.

Cue	Direction	χ^2-value	p-value
Gaze	Towards	0.22	0.64
Point	Towards	49.24	<0.0001
Gaze	Away	0.21	0.64
Point	Away	46.11	<0.0001

Paired comparisons between the two body parts for both cue directions showed no difference between dwell time on the bodies and heads ($\chi^2(1) = 3.42, p = 0.06$).

2.3.2.10 Dwell Times on the target for cues in competition

Dwell times on the target, can be compared without problems as ROI size and location for the different targets was kept constant across the different cueing combinations (including the no cue scene). Figure 2.18 suggests that dwell times on the target are longest for the no-cue condition and shortest for cue combinations not involving an actor pointing towards the target. Similar to the previous analysis on the target for congruent cues, dwell times for this analysis are are expressed as a percentage of the total trial time.

That the seven cue conditions led to different dwell times on the target, was confirmed by a linear mixed-effects analysis ($\chi^2(6) = 150.16, p < 0.0001$). To examine the relative effects of cues pointing towards and cues pointing away from the target, a linear mixed-effects analysis of the effects of the cue pointing towards the target on the dwell times on the target was conducted. This analysis showed that these dwell times were different across the three cues and the no-cue condition ($\chi^2(3) = 13.52, p = 0.004$). Comparisons between the target dwell times for towards cues, showed significant difference between the gaze and pointing cue ($\chi^2(1) = 4.42, p = 0.04$), a significant difference between pointing and arrow cue ($\chi^2(1) = 5.46, p = 0.02$), but no significant difference between arrow and gaze cues

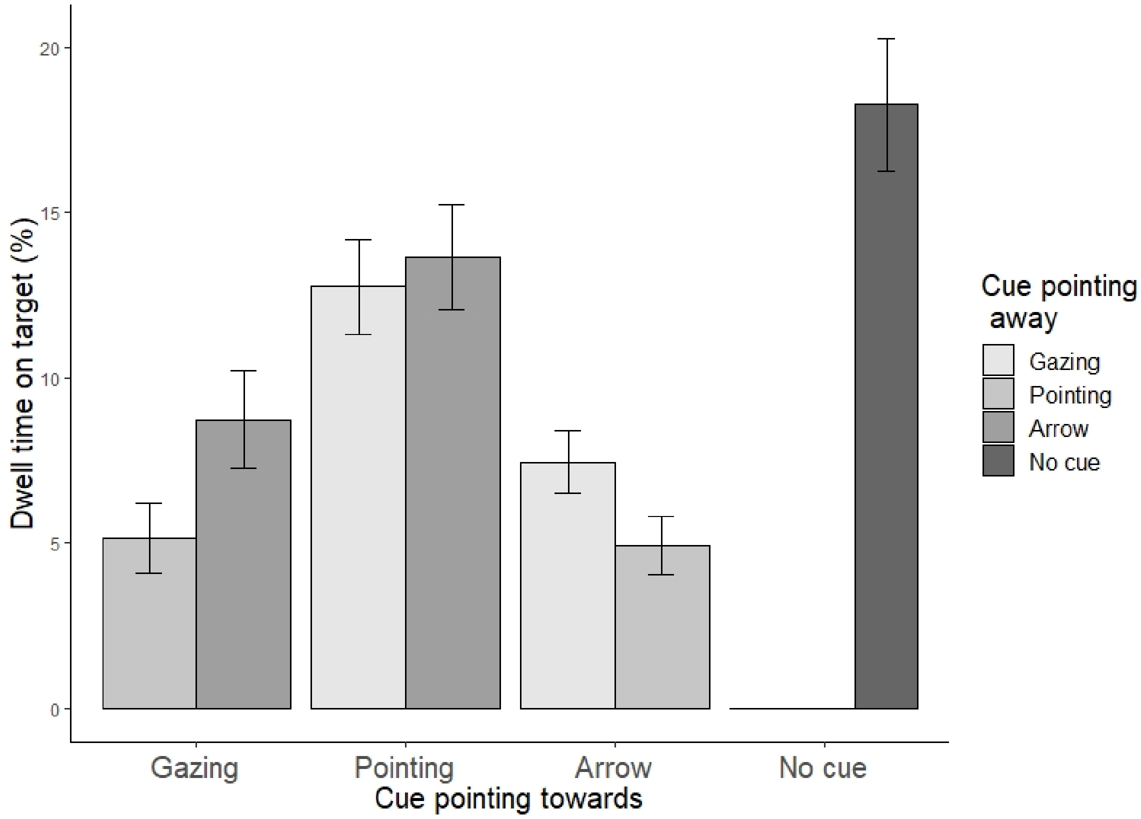

Fig. 2.18 Dwell Times on target for cues in competition. Error bars represent the standard error of the mean across participants.

($\chi^2(1) = 0.08, p = 0.78$). Compared to the no cue condition, dwell times on the target in the absence of a cue were significantly longer. A possible reason is that participants are not distracted away from the target by a cue in this condition. When one of the pointing cues was pointing at the target, dwell times on the target were longer than for the other two cues. This effect is different from the findings in Experiment 1 and the results in the congruent cue conditions of Experiment 2. To compare the congruent and competing cue conditions of Experiment 2 directly, a linear mixed-effects analysis testing the interaction between the towards to the target cue and the direction of the away cue (second cue), showed no significant difference ($\chi^2(1) = 8.04, p = 0.98$). Similarly, there was no significant main effect of the type of target for the towards to target cues ($\chi^2(2) = 4.03, p = 0.13$) and on the targets for the second (pointing away) cues ($\chi^2(2) = 1.53, p = 0.47$). These results suggest that the effect towards to target cue had on dwell times on the target was independent of the cue that it was combined with. Maybe these conflicting results for the arrows is related to the fact that is was accompanied by an informative to the target social cue. This might have led to observers to spend more time on the social cues and to the direction they were pointing at.

2.3.2.11 Trials with fixations on the target in cues with competition

The percentage of trials with a fixation on the target (Figure 2.19) shows a similar pattern as the dwell times on the target earlier, with a significant difference between the cue combinations ($\chi^2(6) = 148.13, p < 0.0001$). As for dwell times, the target was looked at most often if no cue was present. Comparisons between each cue combination can be found in Appendix A.

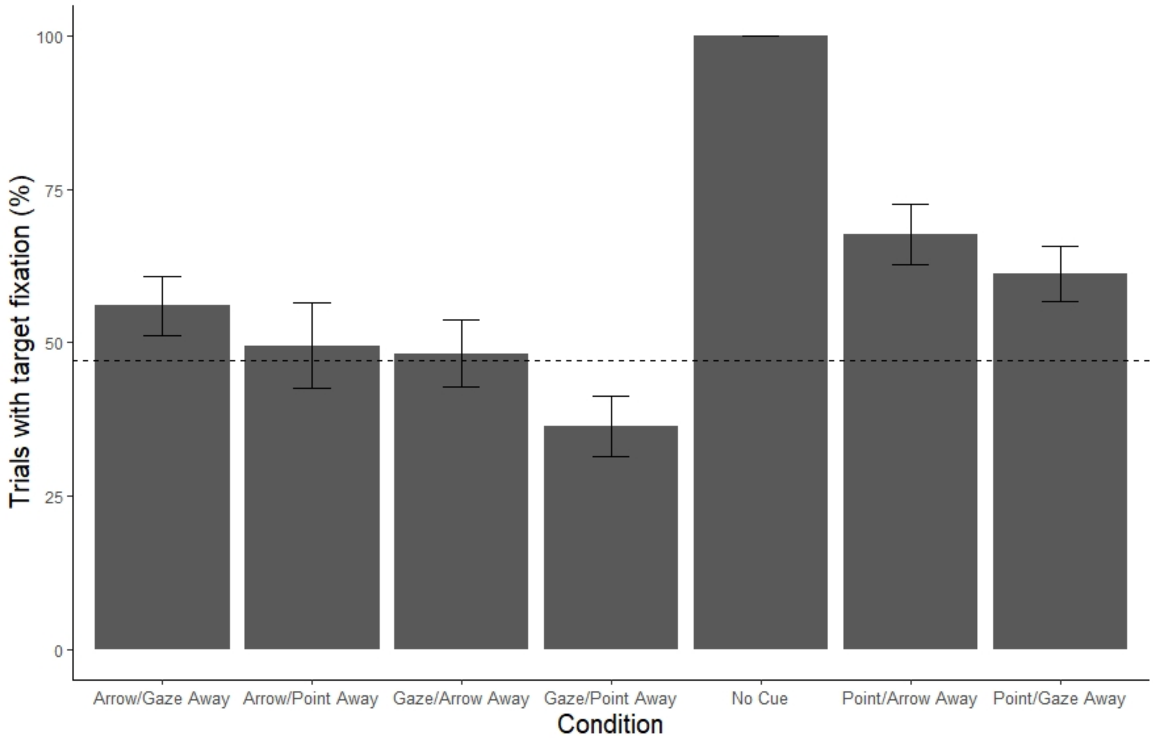

Fig. 2.19 Percentage of trials with at least one fixation on the target. Error bars represent the standard error of the mean across participants.

2.3.2.12 Direction of saccades for cues in competition

Figure 2.20 shows that particularly for gaze and pointing cue most saccades leave the cue, but do not go to the target. The graph also suggests that arrows most often led to a fixation on the target. This result was independent of the cue that the arrow was combined with.

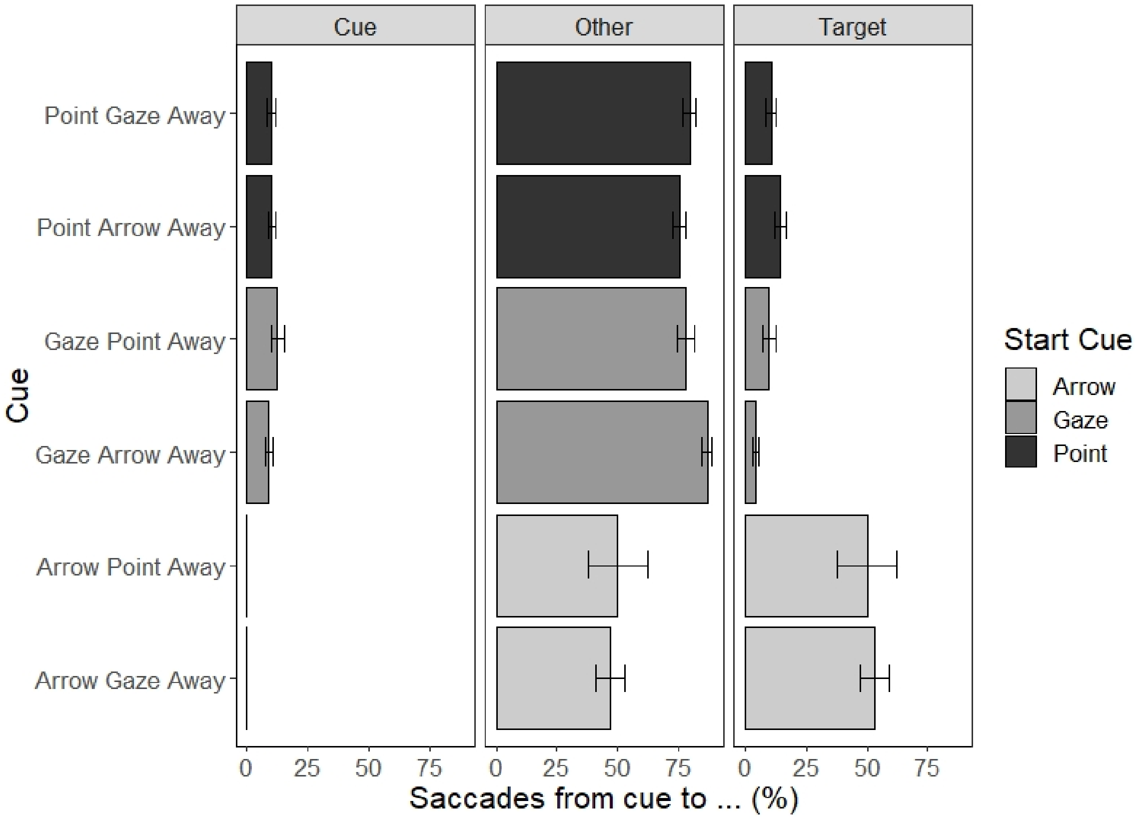

Fig. 2.20 Saccades on the three different directions for in competition cues. Error bars represent the standard error of the mean across participants.

A linear mixed-effects analysis comparing the three saccade directions (cue, target, elsewhere) showed that significantly more saccades went to 'elsewhere' ($\chi^2(5) = 18.57, p = 0.002$). Similar significant differences were found when only the first block of the trials was analysed ($\chi^2(5) = 19.85, p = 0.001$). A linear mixed-effects analysis also showed that the percentage of saccades to the target, depended on the cue combination ($\chi^2(2) = 20.62, p < 0.0001$). First block of trials analysis showed similar results ($\chi^2(2) = 8.12, p = 0.02$). Direct comparison between cue combinations showed the saccades leading to the target were more frequent for arrows pointing to the target than social cues (arrow vs gaze: $\chi^2(1) = 13.83, p = 0.0002$; arrow vs point: $\chi^2(1) = 10.33, p = 0.001$), whereas no difference in the frequency of these saccades was found between gazing and pointing cue ($\chi^2(1) = 1.23, p = 0.27$). When only the first block of trials was analysed, similar significant results were found for the saccades leading the target for arrows when compared with the two social cues (arrow vs gazing: $\chi^2(1) = 7.20, p = 0.01$; arrow vs pointing:$\chi^2(1) = 6.67, p = 0.001$) but different results for the difference between the two social cues ($\chi^2(1) = 4.82, p = 0.03$) with pointing leading more frequent to the target compared to gazing cue. A linear mixed-effects analysis shows that the direction of the saccades depends on the cue combination ($\chi^2(6) = 159.43, p < 0.0001$) with toward arrows leading to more saccades to the target compared to the pointing and gazing cues. Same results were found when only the first block of trials was analysed. Subsequent analysis of the saccades that do not lead to the target but elsewhere showed a significant difference between the cues ($\chi^2(2) = 12.29, p = 0.0002$). First block of trials analysis showed similar significant results ($\chi^2(2) = 5.66, p = 0.04$). Paired comparisons showed no significant differences between gazing and pointing cues ($\chi^2(1) = 0.57, p = 0.45$), but significant differences between arrow and pointing cues ($\chi^2(1) = 6.16, p = 0.01$) and between arrow and gazing cues ($\chi^2(1) = 10.20, p = 0.001$). Similar results were found when only the first block of trials was analysed (gazing vs pointing:$\chi^2(1) = 2.51, p = 0.11$; arrow vs gazing:$\chi^2(1) = 5.04, p = 0.02$ and arrow vs pointing:$\chi^2(1) = 1.94, p = 0.03$). These results suggest that arrows led less to a non-target location compared to gazing and pointing.

2.3.3 Discussion

Past studies on social attention have focused on situations where a single cue is present on the display. At least one attempt has been made to examine how social cues and arrows compete for attention when both present in a visual scene (Birmingham et al., 2009). This study, however, used a very limited number of images, and the cues in the scenes were not matched for saliency, position and size. Experiment 2 tested the effects of two simultaneously presented cues on an observer's attention in a systematic way. Three types of cues were used: Arrows, an actor gazing, and an actor pointing. Cues could either both point (or look) at the target, or one cue could point (or look) elsewhere. Three questions were addressed: (1) which type of cue will capture participants' attention most strongly? (2) which cue shifted attention most strongly towards the target? (3) does the direction of the cues matter?

To address the first question, the time participants spent looking at the cues for the two different cue combinations (congruent and competition) was compared. Findings were in agreement with Birmingham et al. (2009), who found that when a social cue was shown together with an arrow, participants spend most time looking at the social cue. The present results show that this finding is also obtained when the cues are placed in the scene, rather than already present (e.g., arrows painted on the street) and when one of the cues points away from the target object in the scene. One finding that was consistent in both cue directions was that participants showed a clear interest for the people in the scene when the latter were compared with an arrow. The preference of the social cue is similar to both combinations, with pointing being a "worthy opponent" to gazing cues as they were considered equal or more important from gazing

This preference towards the pointing, might be due to various reasons. It can be assumed that dwell time findings might be related to the pointing cues and not to the direction they were presented. More precisely to the meaning pointing cues have. Pointing is a gesture related to directing to a certain location or item (e.g., Langton & Bruce, 2000; Langton et al., 1996). One may wonder if the reason participants preferred an actor that was pointing, over an actor who was gazing (even when it was uninformative), was that they wanted to discover where pointing actor was directing their hand at. Potentially, pointing is considered as a more communicative act, with a message aimed at the observer. On the other hand, gazing might imply that actors gazing or turning their heads to a direction, is just for their own purpose, without having the observer in mind. Moreover, gaze cues are considered to have different (and more significant) social importance to pointing cues. In gazing when the eyes or head are presented averted, their direction inform observers that the person is attending to someone or something else in the environment (Emery, 2000). This ability to attend the same object or part of the scene is considered to be crucial for many aspects in our social life (e.g., T. Allison et al., 2000). Pointing can be equally important for the same reason (e.g., directing attention to avoid a hazardous situation; pointing at a person).

For the gazing cues, it is unquestionable that it is an important cue that still bias people attention towards them. Although when presented alongside with pointing, gazing was not the main cue participants spent time on, it was always noticed, showed a stronger effect from the arrow and captured participants' attention even when uninformative to targets location. For the latter it is not surprising to see that, as in the classic studies (e.g., Driver et al., 1999; Friesen & Kingstone, 1998; Friesen et al., 2004; Hayward & Ristic, 2018; Langton & Bruce, 1999), not-predictive gaze cues captured attention the most compared to informative symbolic cues.

In the analysis, the gaze and pointing cues were defined as the entire body of the actors. Therefore, it is interesting to explore how much time the observer spent looking at the difference body parts (head, body and arm). Analysis on the three body parts could provide a better understanding on how participants distribute their attention. On average dwell times were longest on the heads of the actors, followed by their bodies and their hand (for pointing cues), similar to Experiment 1 where only one cue was present in the scene. These results agree with Perrett et al. (1992)'s model of social attention. Where it is assumed that at the top of the hierarchy of social attention cues, stands the gaze direction, followed by the head direction and body direction. Similarly, strong focus on the head of the actors was also observed in studies examining gaze patterns for natural scenes with one or more people (e.g., Birmingham et al., 2009; Hermens & Walker, 2015).

One important outcome of the dwell times on the three body parts, is that when social cues provided by an actual person, observers do not spend time only on the pointing hand (pointing cues) or head (gazing cues). Instead significant time is spent on the rest of the body. This is a consistent finding with Experiment 1 and it seems that is one important suggestion on how we define social cues when the latter are presented in natural scenes (i.e., considering or not the entire body as a part of the cues). In real-world settings and natural scene, when encountering a person what is considered as a social cue might be more complicated (e.g., more body parts are involved) from the classic approach of just a face or hand. For example, in real life situations when we see someone pointing, we can retrieve information not just from the pointing hand but also by the direction of the head and the body.

When it comes to the arrows in the scene, dwell time results match to the ones in Experiment 1 were a single arrow was presented in the scene. These findings for the arrows are again in agreement with other studies using natural scenes (Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010). Where arrows, alone or together with a gaze cue, failed to capture people's attention. Although it was argued in the introduction (see aims) that arrows in Birmingham et al. (2009)'s study were not given a fair chance against the gaze cue, it seems that this was not the reason for them to looked at less. Moreover, one further aim of this experiment was to see if the direction of the cues will influence the time the observers dwelled on them. This indeed seems to be the case: Dwell times on the arrows were dependent on the direction of the arrow (away and towards the target) and the other cue in the scene (gazing or pointing). To an extent this contradicts previous findings where predictive and counter predictive arrows and social cues had similar effects (e.g., Guzzon, Brignani, Miniussi, & Marzi, 2010; Hermens & Walker, 2010; Kuhn & Kingstone, 2009; Tipples, 2008).

A question that still stands is why observers appear to have very little interest on the arrows (present work and Birmingham et al., 2009). One possible reason is that arrows have less social relevance and their use is restricted to task relevant situations where people search or require directions (e.g., Birmingham et al., 2009; Hayward & Ristic, 2018) (e.g., fire exit signs helping to locate the exit door) . Indeed, if we consider that the majority of the studies exploring the effects of an arrow are in the framework of a task where cues are given a directional meaning, whereas this study (and Experiment 1) no instructions or meaning were assigned to these cues. It therefore seems that if the task is simply to look at an image of a natural scene, observers seek out people presented in the scene. Arrows are considered less important and therefore ignored (Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010). An alternative explanation for the low dwell times on the arrows, is that arrows are more visible in the peripheral vision due to their distinct shape (e.g., Hermens, 2017). However, the current results are not clear, and cannot conclude if the peripheral vision played a role for the arrows results in this experiment and in Experiment 1. It can only be speculated and assumed that this might be one of the explanations. Future studies could explicitly address this question, for example using the moving window paradigm often used in reading research (e.g., Rayner, Pollatsek, Ashby, & Clifton Jr, 2012).

Similar to Experiment 1, dwell times on the cue and targets are not informative on how strongly cues shift attention to the targeted item. Inspired by Hermens and Walker (2015), the present analysis therefore also focused on saccades leaving the cue, once the cue was fixated. Analysis of these saccades suggested that when both cues were pointing at the target, arrows led to more subsequent fixations on the target than gazing cues (see also Hermens & Walker, 2015). As was argued in Experiment 1 (see discussion on saccade) this might be dependent on the meaning gaze cues have for the observer (e.g., spend time to understand actor's emotional state). An alternative explanation might be that gaze cues have been found to only lead to strong cueing when presented at fixation, as discussed in the general introduction (e.g., Hermens et al., 2015). When away from fixation, there are few studies which can support the failure of gaze cues to produce a successful cueing effect, when compared with an arrow or another a pointing cue (e.g., Burton et al., 2009; Hermens et al., 2015).

Pointing cues provided a strong competition with the arrows in terms of saccades when both cues pointed at the target. At least two explanations can be offered for these findings. First, both cues communicate a clear directional message. Pointing and arrows are meant to point at something, in contrast to gaze cues, which can also occur without the aim of directing another person's attention to an object. Second, both extrafoveally presented arrows and pointing hands are considered to possess a more distinct shape compared to gazing cues, which allows them to stand out from the environment (e.g., Hermens, 2017). Therefore, when these cues are presented together and pointing at the same direction, they might be equally processed and lead similarly to the target.

Interestingly, saccades of gaze and pointing cues showed similar cueing effects when their direction was towards the target. There is a limited number of studies which have presented both pointing and gazing cues at same time (e.g., Langton & Bruce, 2000; Langton et al., 1996). Using a Stroop interference task, Langton and Bruce (2000), presented pointing and gazing cues simultaneously (provided by the same person). Cues were pointing either at the same or the opposite direction. Langton and Bruce (2000) suggested that when pointing and gazing have the same direction, information for the location of the target is processed in parallel. Observers showed an automatic extraction of directional information from both cues. Based on Langton and Bruce (2000)'s findings, it can be assumed that when pointing at the same target, observers in the current experiment used both social cues to retrieve information about the targets' location. Based on this assumption, an alternative explanation can be offered. When presented together, pointing cues might enhance or give a clearer directional meaning to the gazing cues and vice versa. This can be reflected by comparing the percentages of successful shifting towards the target, between competitive and congruent gaze and pointing cues. When only one social cue was pointing towards the target, cueing shifting effect was fairly low for both cues (around 10%). Whereas, when pointing and gazing cues both directed observers' attention to the same location, success rates were overall increased (around 22%).

As in the other two experiments, saccades were substantially more often directed to a position other than the target than at the target. One possible reason is that that the direction of the cues can be ambiguous (although an effort had been made to photograph the cues from the side to avoid spatial distortions). Another factor is that saccades could not only be directed to the target, but also to objects in between the cue and the target (although effort had been made not to have salient objects between the cue and the target). Both types of effects do not occur in more traditional studies of cueing effects, as there is little else in the display to look at than the target and the cue (Friesen et al., 2004; Kuhn & Kingstone, 2009). This shows again that it is important for cues to be shown in their natural environment, as restricting the number of objects in on the screen may lead to artificial gaze behaviour not seen in normal day-to-day vision. The limited number of saccades to the target agrees with Experiment 1 and Hermens and Walker (2015).

Overall, the results from Experiments 1 and 2 largely agree. Social cues attract more attention than arrows, but arrows and pointing cues lead to more saccades to the target. The two experiment so far, however, have asked participants to 'simply look', which may not be a common task, as most viewing is done with a purpose (e.g., to locate an object). Therefore, the last experiment of this chapter will examine how social and symbolic cues affect the observers' attention under a visual search task.

2.4 Experiment 3

Experiment 1 and 2 explored whether social and symbolic cues affect people's attention when a single or multiple cue are presented in a free viewing task. Experiment 3's purpose is to investigate whether dwell times on the cues depend on the task that people are performing.

Experiment 3 therefore asked participants to perform a visual search for a target object in the scene, presented as a small image before the natural scene containing the arrows or actors.

2.4.1 Methods

2.4.1.1 Participants

Forty-one students (Males = 11, Females = 30) from the University of Lincoln were recruited for this study. Their age ranged from 18 to 24 (Mean = 22.20, SD = 6.06) and all of them reported normal or corrected to normal vision. All participants provided a written consent form for the study approved by the local ethics committee and received course credit in return for their participation.

2.4.1.2 Apparatus

As in Experiment 1 and 2, stimuli were presented on a 19-inch Dell flat screen (with resolution 1,280 x 1,024; approximately 26° x 20° of visual angle) using the Experiment Builder (SR Research, ON, Canada) software. Participants were positioned 80 cm away from the screen (distance from the screen was controlled with a chin rest). Participants' eye movements were recorded with an Eyelink 1000 desk-mount eye tracker, sampled at a 1000Hz rate. A standard USB mouse was used to collect participants' responses.

2.4.1.3 Stimuli

Twenty-one scenes showing indoors and outdoors areas from the University of Lincoln campus were used as the stimuli scaled down to the size of 800 x 600 pixels (approximately 16° x 12° of visual angle) using the IrfranView software and presented on a 1,280 x 1,024 pixels display. Images were scaled down for the same reasons described in Experiment 1 (see section 2.2.1.3). Each scene was photographed four times, for each of the four conditions (pointing, gazing, arrow, non-cues). To create the target images, presented before the natural scenes, the areas around two items within the photos were cropped from the scenes with the IrfranView software. The cues' location was the same in all versions of each scene. Across scenes their position varied according to items location (Figure 2.21).

Fig. 2.21 Example of stimuli showing the two different versions of gazing cues presented to the participants (different groups). The left image shows the actor looking at the brown package on the table. The right image shows the same actor looking at the opposite direction towards the blue sign.

2.4.1.4 Design

Experiment 3 had a mixed measure design. All participants performed a visual search task making the design of this experiment a repeated measures, ensuring that all scenes were seen by each participants. Participants were assigned to four different groups, seeing different combinations of informative and uninformative cues, making the design a between subject design. The total list of scenes and cues was divided into four blocks where similar number of cues were presented. The order of the blocks was counterbalanced across participants and trials within each block were randomised for each participant.

2.4.1.5 Procedure

At the beginning of the experiment participants were asked about their handedness. Left-handers were asked to use the mouse with their left hand. Written instructions were provided, and participants signed the consent form. They were then asked to place their head in the chin rest and their hand on the mouse, after which a 9-point calibration was performed. Each trial started with a drift correction and confirmation of fixation by the experimenter. A picture of the target object was then shown for 1000 ms followed by a photo containing that target for 5000 ms. Participants were asked to click on the target as quickly as possible. After each 21st trials participants were offered a short break and a re-calibration was performed where needed. No feedback was provided on whether participants correctly identified the target. After completing the total of 63 trials, participants were debriefed about the purpose of this study and were given the chance to ask questions about the experiment. Participants were also asked a few questions (e.g., if they noticed the cues, if they were aware that the cues' orientation was leading to the target object on some of the trials). This information were neither recorded nor analysed. The aim of the questions was to engage with the participants and ensure they understood the purpose of the experiment.

2.4.1.6 Data Analysis

Before analysing the results, trials with incorrect responses were removed. Fixations were then assigned to regions of interest using a method similar to Experiments 1 and 2. Data visualizations and statistical analyses were conducted with the R software (version 3.4.1, R Development Core Team 2017). During exploration of the findings, it turned out that the Experiment Builder software, for unknown reasons, had not collected all participants' responses for some of the trials. This resulted in dropping 40% of trials for the arrows and 36% from each social cue. While these missing trials pose an issue for tests that expect a complete design (ANOVA, t-test), no such issues with analysis and interpretation of the data are expected with linear mixed-effects models. This analysis deals with missing data of this kind by incorporating random effects for the stimuli, meaning that differences between trials are compensated for in the analysis. Therefore, for this experiment only linear mixed-effects models, with Bonferroni correction, will be used to analyse these data. Similar to Experiment

1 and 2, first block of trials analysis will also be reported for the dwell times on the cues and saccades analyses to explore any effect from the repetition of scenes in the results.

2.4.2 Results

2.4.2.1 Reaction Times and cues

Before reactions time were computed, outliers (reaction time was higher than 3000ms as well as incorrect responses (not clicking the correct item) were removed from the analysis. Reactions times for each condition per validity were calculated. Figure 2.22 shows the average correct reaction time for cues pointing at the target (predictive) and those pointing in a different direction (no predictive), for the three cueing conditions. Suggest that the reaction times for the three cues did not significantly differ and this was consistent for both predictive and not-predictive to target location conditions.

A linear mixed-effects model analysis showed an interaction for reaction times between the cueing condition and the two targets (pointed-at or elsewhere) ($\chi^2(6) = 42.43, p < 0.0001$). A follow up analysis compared the effect of validity per cue type using a linear mixed-effects analysis, showing an advantage for arrows ($\chi^2(2) = 70.72, p < 0.0001$) compared to the gaze and pointing conditions. No differences in reactions times were found between informative gaze and pointing cues and uninformative gaze and pointing cues ($p > 0.05$). For each of the condition, there was no significant difference between informative ($\chi^2(3) = 3.82, p = 0.28$) and uninformative ($\chi^2(3) = 4.56, p = 0.21$) cues. This suggests that the mere presence of the cue helped finding the target, but the direction of the cue did not matter. One potential reason may be the influence of prior knowledge about where the target could be found in the scene (e.g., standing objects tend not to be near the ceiling). In addition, these results need to be treated with caution, because the dropped trials may have had an effect on the results. A follow up experiment should provide clarity on whether the dropping of trials had an effect. While reaction times are the most important measure of

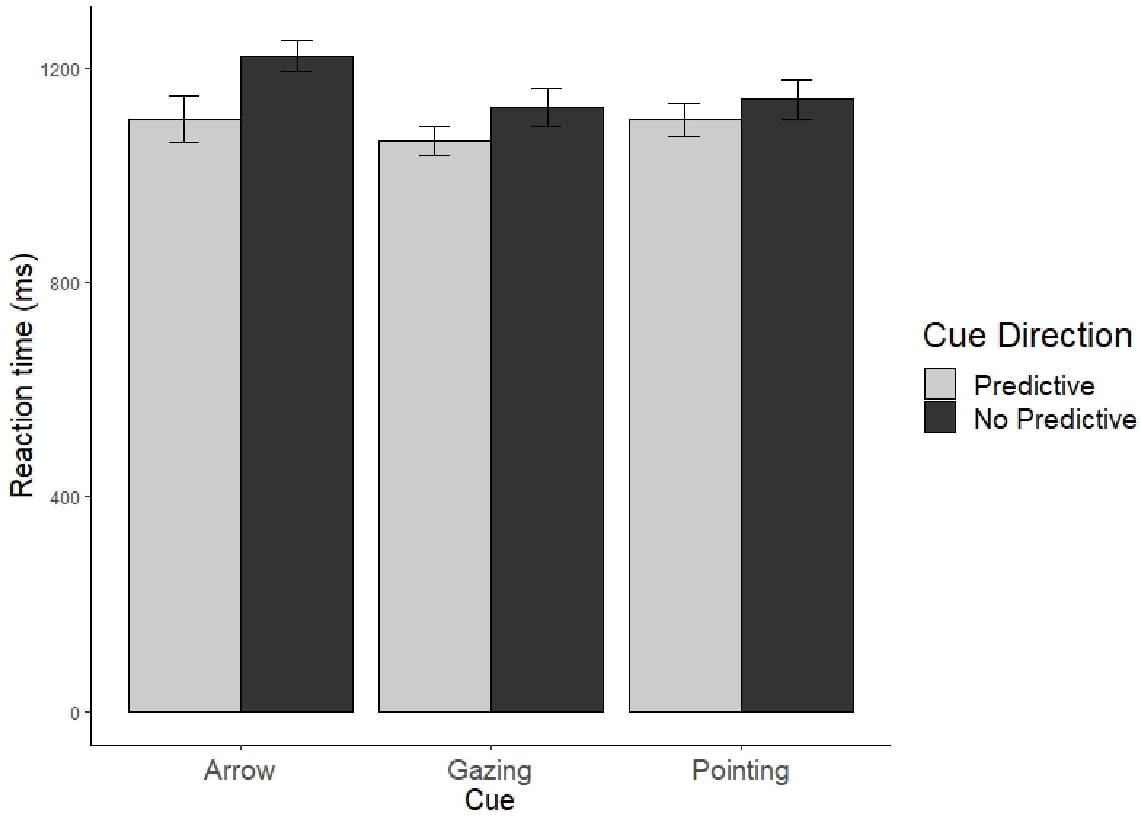

Fig. 2.22 Reaction Times of the three cueing conditions and the two cues' directions (predictive means that the cue was pointing at the target; no predictive means that the cue was pointing away from the target). Error bars represent the standard error of the mean across participants.

covert attention, eye tracking provides an excellent measure of overt attention, which tends to be the more prominent way of shifting attention. A subsequent analysis will therefore look at participants' gaze behaviour to explore these questions.

2.4.2.2 Dwell Times on the cues

We here use the same measures as in the first two experiments, and also start with including the body in the regions of interest of gazing and pointing cues. Figure 2.23 plots the dwell times (expressed as a percentage of the total trial time) on the cues separately for the two different cue direction. The results suggest that dwell times on the cues were longer for gaze and pointing cues than the arrows. Note, however, that dwell times on cues were very short in comparison with those during free viewing in Experiments 1 and 2. Note also that the 'no cue' condition is not included in this graph, as there is no cue to look at.

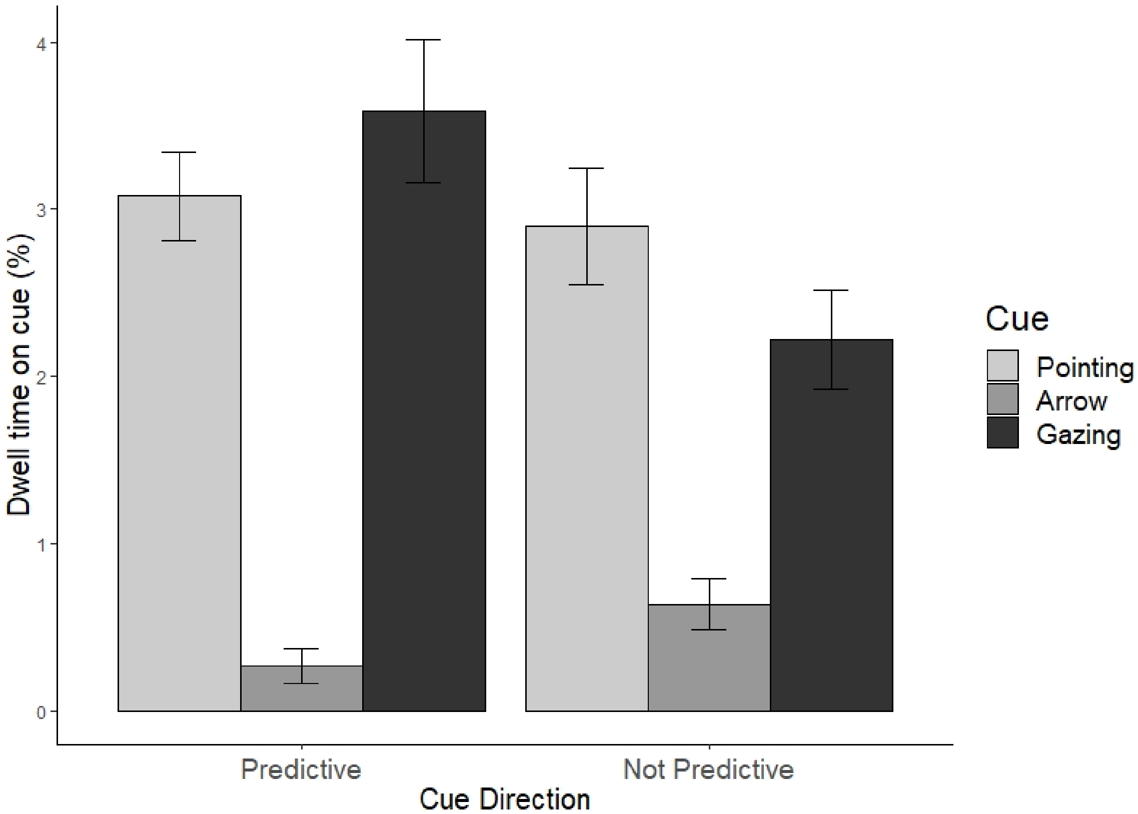

Fig. 2.23 Dwell Times on the cues per cue direction. Predictive direction indicates that the cue was informative to targets' location. No Predictive indicates that the cues direction was not informative to targets' location. Error bars represent the standard error of the mean across participants.

This difference between dwell times on the gaze and pointing cues and the arrow cues was confirmed by a linear mixed-effects analysis, using participants and scenes as random factors, comparing a model with and without the condition factor ($\chi^2(3) = 243.43, p < 0.0001$). First block of trials analysis showed similar significant results ($\chi^2(3) = 134.53, p < 0.0001$). Linear mixed-effects paired comparisons between the three cue conditions were computed, showing no significant differences between gazing and pointing conditions ($\chi^2(1) = 0.06, p = 0.80$), significant differences between pointing and arrow conditions ($\chi^2(1) = 98.25, p < 0.0001$) and significant differences between gazing and arrow conditions ($\chi^2(1) = 94.74, p <$

0.0001). Similar results were found when only the first block of trials was analysed (gazing vs pointing:$\chi^2(1) = 0.25, p = 0.92$; gazing vs arrows:$\chi^2(1) = 32.44, p = 0.001$ and arrow vs pointing:$\chi^2(1) = 42.24, p < 0.0001$).

The statistical analysis thus far has ignored the direction of the cues (pointing towards or away from the target object). To explore the effect of cue's direction an interaction analysis was conducted between the three cues and two cues' direction (Predictive; Not-Predictive). Dwell times on the cues for the two different conditions show that, in general, cues were affected by whether or not the cues pointed at the target (linear mixed-effects analysis of the interaction between cue type and cue direction ($\chi^2(5) = 106.15, p < 0.0001$). First block of trials analysis confirmed this significant result ($\chi^2(5) = 62.07, p < 0.0001$). Linear mixed-effects pairwise comparisons per cue type showed that only for gaze cues, cues that pointed at the target were looked at for longer ($\chi^2(1) = 7.93, p = 0.005$). Same difference was found when only the first block of trials was analysed ($\chi^2(1) = 4.75, p = 0.02$). Dwell times also differed between cues, both when the cues pointed at the target ($\chi^2(6) = 32.64, p < 0.0001$) and when the cues pointed away from the target ($\chi^2(6) = 69.13, p < 0.0001$). First block of trials analyses confirmed these results. The significant differences between predictive cues were due to a significant difference between the arrow and the pointing cues ($\chi^2(5) = 83.11, p < 0.0001$) and between the arrow and the gazing cues ($\chi^2(5) = 75.68, p < 0.001$. No difference was found in dwell times on predictive gaze and pointing cues ($\chi^2(5) = 0.98, p = 0.32$). These difference between the predictive cues were also confirmed by analysing the first block of trials (arrow vs gazing:$\chi^2(5) = 24.11, p = 0.01$; arrow vs pointing:$\chi^2(5) = 32.91, p = 0.002$ and gazing vs pointing:$\chi^2(5) = 2.06, p = 0.61$). Similar results were found for the not predictive cues, with a significant difference between the arrow and the pointing cues ($\chi^2(5) = 59.89, p < 0.0001$) and a significant difference between the arrow and the gazing cues ($\chi^2(5) = 66.718, p < 0.0001$), but not between the gazing and pointing cues ($\chi^2(5) = 2.12, p = 0.15$). These results were also found when only the first block of trials was analysed (arrow vs pointing:$\chi^2(5) = 28.66, p < 0.0001$; arrow vs gazing:$\chi^2(5) = 15.89, p = 0.007$ and gazing vs pointing:$\chi^2(5) = 1.22, p = 0.29$).

2.4.2.3 Trials with fixations on the cue

Figure 2.24 shows the percentage of trials with at least one fixation on the cue or target (pooling data across the two cue directions). This plot suggests that social cues had many more trials with at least one fixation on the cue than arrows (i.e., they were much less commonly ignored).

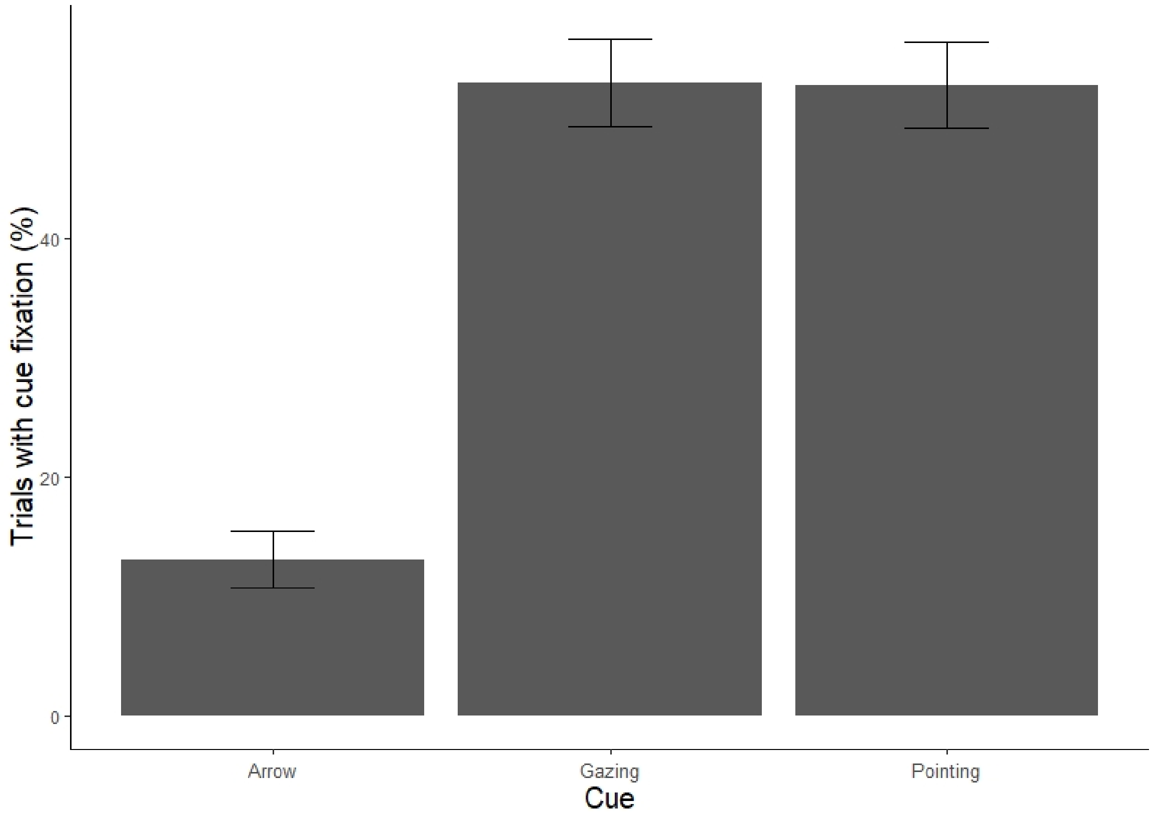

Fig. 2.24 Percentage of trials with at least one fixation on the cue for the three cue conditions (horizontal axis). Error bars show the standard error of the mean across participants.

A linear mixed-effects logistic regression testing these differences in the percentage of trials with at least one fixation on the cue showed a significant effect of cueing condition ($\chi^2(2) = 168.83, p < 0.0001$). Pairwise comparisons showed a lower percentage for arrow cues than for gaze cues ($\chi^2(1) = 10.29, p < 0.0001$) and for arrows than for pointing cues ($\chi^2(1) = 10.26, p < 0.001$). No significant difference was found between gaze and pointing cues ($\chi^2(1) = 0.09, p = 0.76$).

2.4.2.4 Dwell Times on Head, Body and Arm

Figure 2.25 splits the dwell times on the cues in the different sub-regions (body, head, hand), shown separately for cues pointing at the target or away. Similar to Experiment 1 and 2, dwell times on the three body parts for gazing and pointing are calculated on how much time (total) spend on the person (given that they looked at the different body parts).

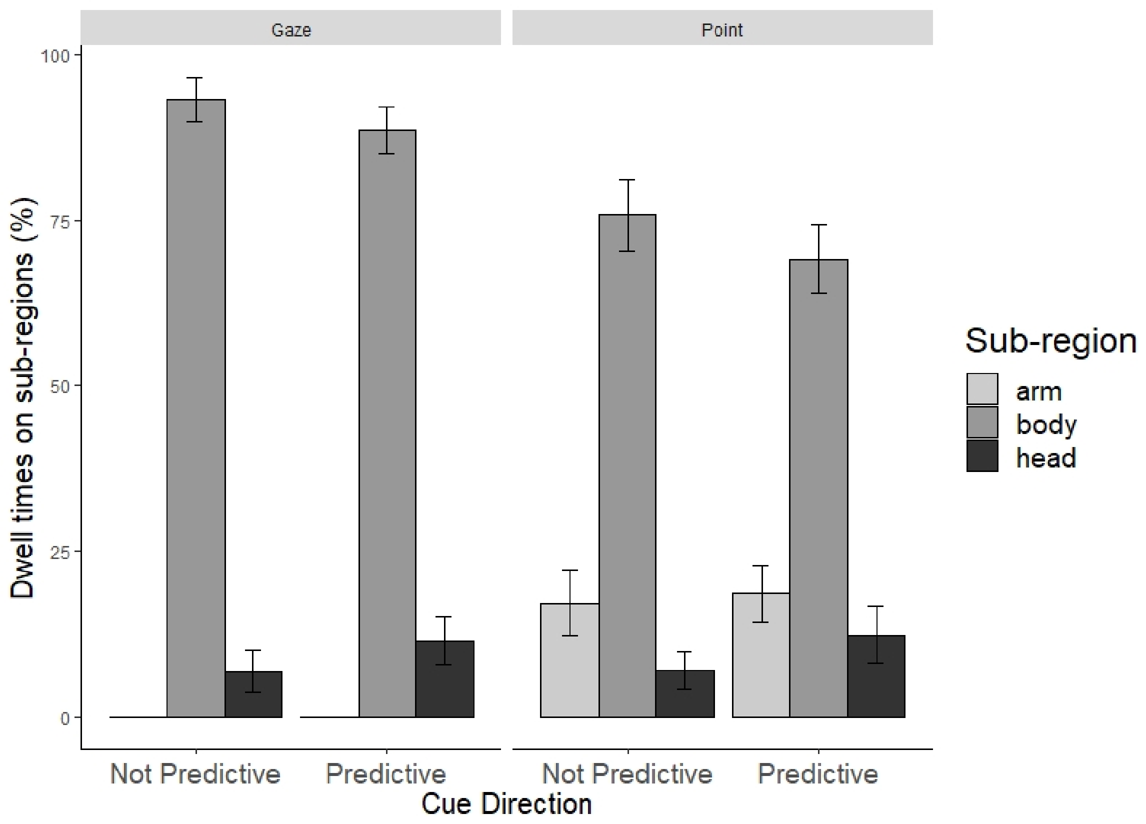

Fig. 2.25 Dwell Times (as a percentage of overall viewing time) on the three body parts (different bars) for the two social cues (gaze cue on the left, pointing cue on the right) and the two cue directions (to the search target = predictive, elsewhere = not-predictive). Error bars represent the standard error of the mean across participants.

Figure 2.25 shows that most of the dwell times on the actor was directed towards the body. A linear mixed-effects model (with participants and scenes as random factors) testing a three-way interaction between the cues, validity and three sub-regions showed not significant difference ($\chi^2(2) = 0.52, p = 0.77$). When arm was excluded from the analysis same results were found ($\chi^2(1) = 0.87, p = 0.35$). Two linear mixed-effects model analyses (one for each cue) tested the interaction between the direction of the cue and the sub-regions (with arm only for pointing cues) to determine whether dwell times were differently distributed depending on whether the cue was pointing at the target or not. No interaction was found (gazing:$\chi^2(2) = 1.04, p = 0.31$ and pointing:$\chi^2(2) = 2.72, p = 0.26$), meaning cue direction did not influence gaze patterns on the cues. Linear mixed-effects analyses of the effects of sub-regions and cue direction showed a significant difference between dwell times on the arm, head and body for pointing (predictive:$\chi^2(2) = 85.84, p < 0.0001$ and not-predictive: $\chi^2(2) = 115.4, p < 0.0001$) and significant difference between dwell times on the head and body for gazing cues (predictive:$\chi^2(2) = 183.87, p < 0.0001$ and not-predictive:$\chi^2(2) = 168.47, p < 0.001$).

As these results differ from Experiment 1 and 2, an additional analysis was conducted to explore whether the dwell times on the body were dependent on the position of the actor in the scene (closer to the center of the scene or more to the periphery). Seven out of the 21 scenes were depicting actor's position closer to the center whereas in fourteen scenes actor's position was more to the periphery. Results for these two analyses can be found in Appendix A. Both analyses agree that actors' position in the scene did not influence the dwell times on the body and that as in initial analysis (all scenes together) participants spent more time looking at actors' body.

2.4.2.5 Dwell Times on the target

As in Experiments 1 and 2, there is a possible confound between the ROIs size, location and dwell times on the cues. While dwell times on the target did not have such confounds in Experiments 1 and 2, Experiment 3 has a part confound in that the target for the predictive and not-predictive cue conditions was a different object. This is something to take into account when interpreting the results. Figure 2.26 shows the dwell times on the target for the three different cue types and whether or not the cue was pointing at the target (in the no-cue condition, these were the same objects, but there was no cue). Note that dwell times are expressed as a percentage of the total trial time.

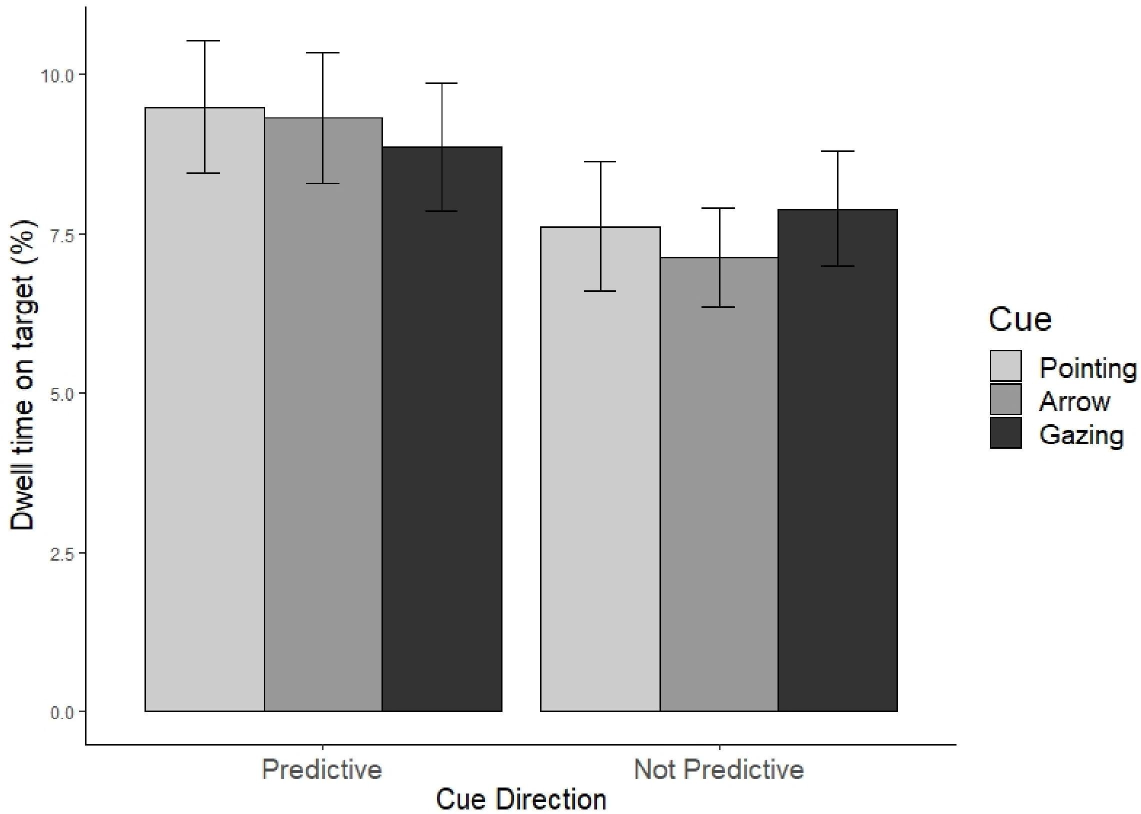

Fig. 2.26 Dwell Times on the target for the three cue types and two cue directions (towards the target and elsewhere). Error bars represent the standard error of the mean across participants.

Dwell times on the target appear to be shorter for the not-predictive condition, but mostly for the conditions with a cue (the black bars, indicating no cue, look more similar). A linear mixed-effects analysis (using participants and scenes as random factors), however, showed no significant interaction between the three cue types and the two conditions ($\chi^2(7) = 10.77, p = 0.15$). A significant main effect of cue direction was found for arrow cues ($\chi^2(1) = 5.27, p = 0.22$) and a marginally significant effect of cue direction as found for the pointing cue ($\chi^2(1) = 3.80, p = 0.51$). No significant main effect of cue type was found (either in the predictive ($\chi^2(3) = 2.03, p = 0.57$, or the not-predictive condition $\chi^2(3) = 0.45, p = 0.93$). Overall dwell times on the target were relatively short, but this could be because participants were exposed to the targets' characteristics (i.e., shape, colour) beforehand, and ended the trial as soon as the target was found.

2.4.2.6 Trials with fixations on the target

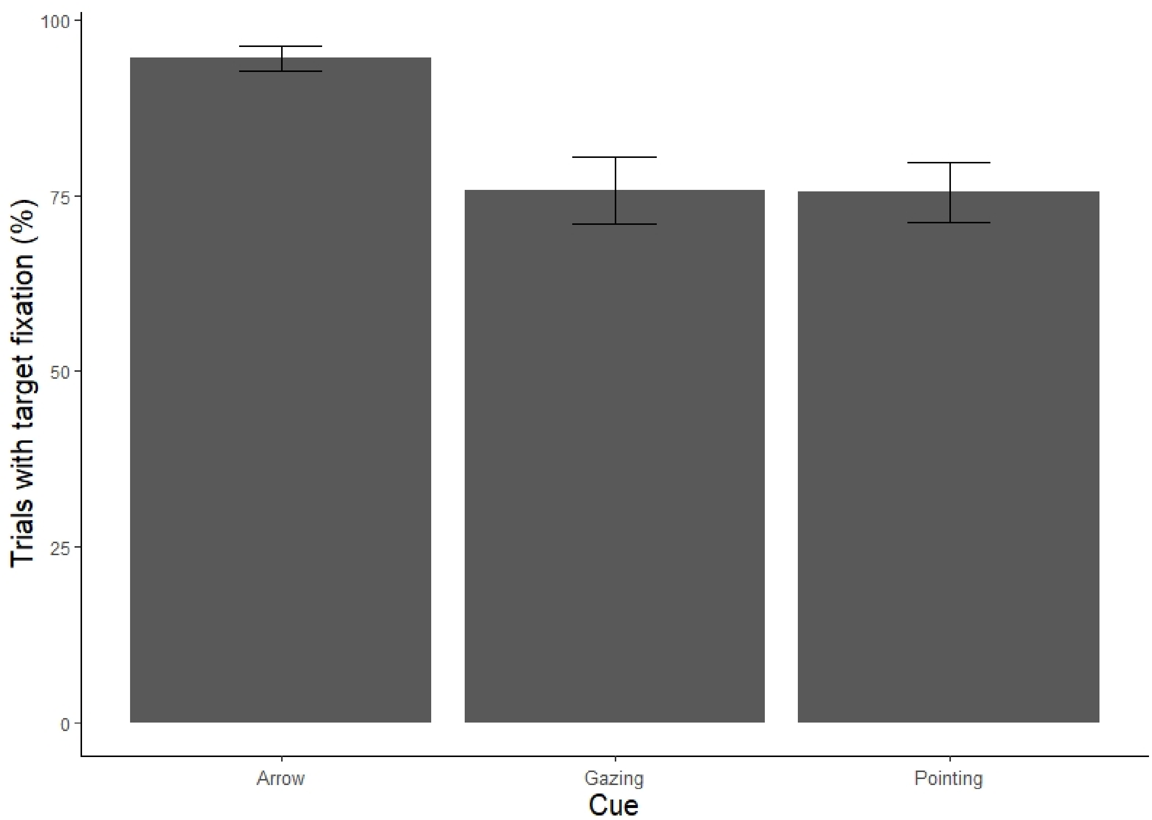

Fig. 2.27 Percentage of trials with at least one fixation on the target for the three cue conditions (horizontal axis). Note that percentages are for both predictive and not-predictive direction of cues. Error bars show the standard error of the mean across participants.

Figure 2.27 shows the percentage of trials with at least one fixation on the target (pooling data across the two cue directions). Dwell times and fixations on cues are somewhat difficult to compare, because visually cues cannot be fully matched in their visual properties. This issue does not arise with fixations on the target (which is always the same object), and differences between cues are much smaller here (see Figure 2.27).

A logistic regression showed a significant effect of cueing condition on the percentage of trials with at least one fixation on the target ($\chi^2(3) = 89.60, p < 0.0001$). Pairwise comparisons showed no significant difference between the gaze and pointing ($\chi^2(1) = 0.10, p = 0.81$) but significant difference between the arrow and the two social cues (arrow vs gaze: $\chi^2(1) = 23.81, p < 0.0001$; arrow vs pointing: $\chi^2(1) = 19.99, p < 0.0001$).

2.4.2.7 Direction of saccades

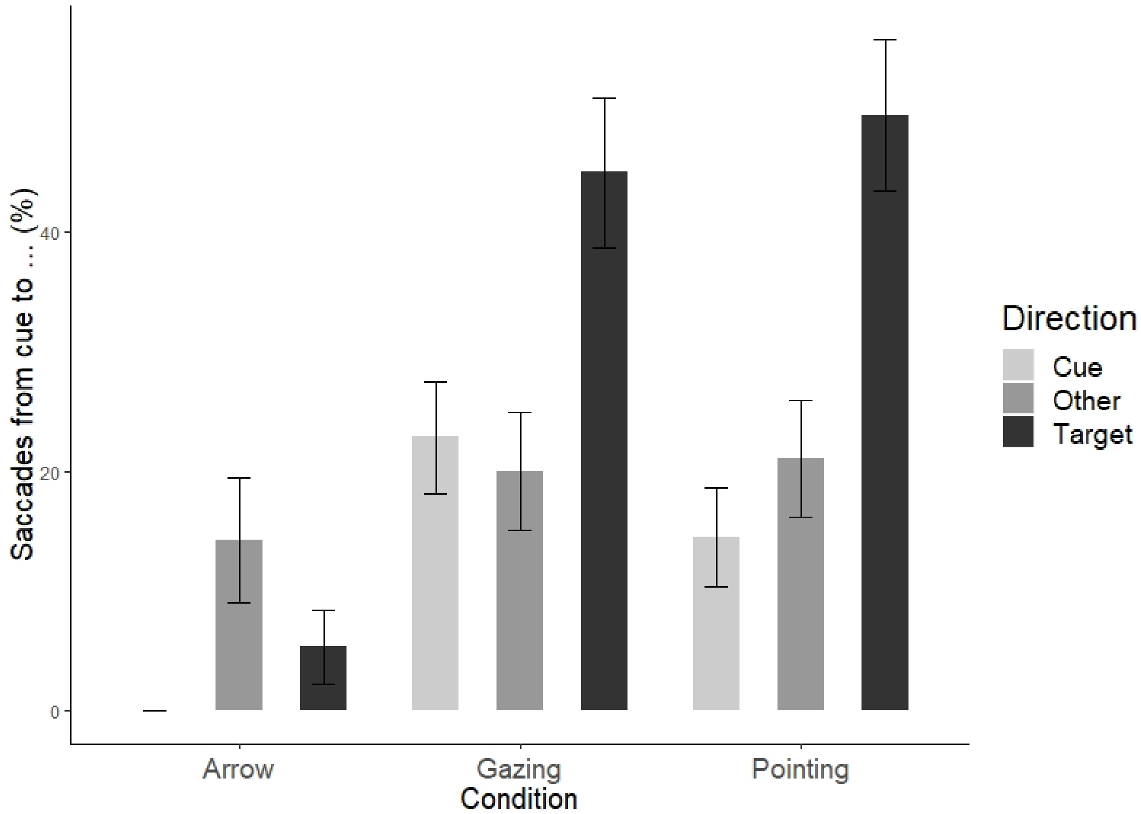

Fig. 2.28 Saccade direction (cue, target, or elsewhere) for the three cue types (arrow, gazing, and pointing cue) when the cue pointed at the target. Error bars represent the standard error of the mean across participants.

The percentage of trials with at least one fixation on the cues is low. Still, it is interesting to explore the direction of saccades for such cue fixations (Figure 2.28).

Figure 2.28 suggests that social cues more often led to immediate fixations on the target, compared to the arrow. A linear mixed-effects model analysis confirmed this observation, showing a significant interaction between the cue type and the direction of the saccade ($\chi^2(2) = 7.03, p = 0.02$). First block of trials analysis showed similar results ($\chi^2(2) = 2.58, p = 0.01$). Linear mixed-effects logistic regression showed no significant interaction between the conditions and the direction away from the target ($\chi^2(2) = 1.14, p = 0.57$). When only the first block of trials was analysed similar results were found ($\chi^2(2) = 4.11, p =$

0.81). Target directed saccades differ between the three cue conditions($\chi^2(2) = 21.46, p < 0.0001$) with pointing and arrows significantly differ ($\chi^2(1) = 20.02, p < 0.0001$), gazing and arrows also significantly differ ($\chi^2(1) = 14.81, p < 0.0001$), but pointing and gazing not significantly differ ($\chi^2(1) = 0.34, p = 0.56$). Similar results were found when only the first block of trials was analysed (target directed saccades:$\chi^2(2) = 11.83, p < 0.0001$; gazing vs arrow:$\chi^2(1) = 5.58, p = 0.002$; pointing vs arrow:$\chi^2(1) = 9.77, p = 0.004$ and gazing vs pointing:$\chi^2(1) = 1.55, p = 0.61$).

2.4.3 Discussion

Past work has suggested that eye movement patterns strongly depend on the task of the observer (e.g., Borji & Itti, 2014; Yarbus, 1967). However, the majority of these studies have merely showed a task effect and not the effect a task can have on the cues. Only recently a very limited number of studies have tried to explore how task (i.e., memory task) affects gaze behaviour on the cues when they are embedded in natural settings (Hermens & Walker, 2015). Experiment 3 was the first to directly compare eye movements of observers when shown natural images with three types of cues (two social and one symbolic) when searching for a target. By using a search task, the effects of the cues on overt (eye movements) and covert (without eye movements) attention could be studied.

The search task has a strong effect on how often and how long the cues were looked at, in line with previous studies that showed strong effects of a viewer's task on their eye movements (e.g., DeAngelus & Pelz, 2009; Fletcher-Watson et al., 2008; Yarbus, 1967). Whereas dwell times on cues were as high as 70% in Experiment 2, dwell times dropped to around 3% when observers performed a search task. Even when cues were fixated, they led to fewer subsequent fixations on the target when a search task was performed. Previous studies (e.g., Dukewich, Klein, & Christie, 2008; Fletcher-Watson et al., 2008; Itier et al., 2007; Tatler et al., 2010; Zwickel & Võ, 2010) have suggested that task can influence observers' eye movements. In addition, due to the methodology of this task (to locate an already known target) there is a possibility that participants were not dependent on the cues to locate the targeted objects but merely used targets characteristics to infer their location in the scene (Malcolm & Henderson, 2009; Zelinsky, 2008). Reaction times and the time participants spent looking the targeted item certainly support such a hypothesis, as there were no differences between cueing conditions on these measures.

The present Experiment is the first to use a visual search task, social and symbolic cues, and natural scenes. Other studies have examined search for more impoverished displays. One study found that when arrows direction points in a different direction, locating a target is more difficult (Jonides, 1981). Other studies found that social cues made it easier to find a search target (Kuhn, Pagano, Maani, & Bunce, 2015; Kuhn & Tipples, 2010). While these studies focused on cues presented at fixation, the present studies cues were often presented away from fixation, where other cueing effects seem to occur. For example, Burton et al. (2009) and Hermens (2017) found that gaze cues were not as effective when presented away from fixation, which agrees with the present findings (although head direction was used here as gaze cue, rather than just gazing eyes).

A question which needs addressing, is how people located the targets if they did not use the cues. A potential mechanism is semantic guidance (Wolfe, Võ, Evans, & Greene, 2011). Semantic guidance is the prior knowledge about items probable location in the given scene (Torralba, Oliva, Castelhano, & Henderson, 2006; Võ, 2010) and the probability that an item will be part of the scene (Võ & Henderson, 2009) (e.g., expecting to see a land mower in the lawn and not in the living room). A further mechanism that may be at play is episodic guidance. Episodic guidance refers to the memory for previously encountered scenes resulting in a memory for information about specific items and locations (Hollingworth, 2006; Wolfe et al., 2011). As all targeted items fitted in each scene's context and scenes were repeated (although not the specific combination of target and scene), these mechanisms may have played a role to targets localization.

Although dwell times on the cues were low, it is worth noticing that both social cues still managed to capture the observer's attention. Interestingly dwell times on social cues did not seem to depend on the direction of the cue (directed at the target or not). If these results are compared with Experiment 1 in this thesis (where only one cue was presented in the scene), dwell times on the cues are contradictive. In experiment 1 gazing was the cue, strongly capturing observers' attention. In the current experiment pointing seems to be equally strong with gazing cues. There is a major difference between the two studies potentially leading to these results. In free viewing (Experiment 1), participants were browsing at an image. Cues were not assigned to a specific role, so the meaning social cues pose in the scene could be left open for interpretation. On the other hand, while performing a search task, observers are looking for a specific item in the scene. As pointing is a directional gesture, when at task it conveyed information relevant to the search task (e.g., help to target localization). Therefore, performing a task, gave a contextual meaning to pointing leading to more dwell times to this cue and potentially having the same semantic meaning with gazing.

The body was found to be looked at most, although it did not contain as much directional information as the hand and head of the actor. This finding was unexpected, Experiments 1, as well as other studies (e.g., Birmingham et al., 2009; Hermens & Walker, 2015; Perrett et al., 1992) had suggested that observers' look more at the heads of the actors. Two explanations can be offered for these findings. First, high dwell times on the body, might be a search strategy due to the task and the restricted available time for completing said task. As participants' goal was to locate a known item as soon as possible, it can be assumed that there was minimum time to spend on the other body parts. From Yarbus (1967) it is well established that when performing a task, eye movement patterns greatly differ from simply viewing a scene. Based on this, it can be assumed, that the difference between this experiment and Experiment 1 of the dwell times on the actors' body regions was due to the influence of the task. Looking more at the body could have been the optimal strategy leaving more search time to locate the targeted item.

One alternative reason for the dwell times on the body, could have been its position in the scene. Although great care was taken to avoid presenting actors close to the middle of the scene, restrictions from the environment (e.g., available space), led to a minority of actors' body to appear closer to the center. When separating the scenes into centrally and peripherally presented bodies, no difference in results was observed. Based on this, it might be that the body was looked at more, due the search strategy discussed above. Although this is not a question which has been studied extensively, there is some recent evidence (e.g., Azarian et al., 2017; Zwickel & Võ, 2010) which have suggested that the body is equally salient to the eyes and head and can produce similar cueing effects.

It could have been expected that arrows would be looked at more in a search task, as arrows clearly signal a direction. This is not what was found. Birmingham et al. (2009) have suggested that the reason arrows fail to capture observers' attention is that these cues are restricted to their cueing effect and nothing more. Consistent with the previous findings in Experiment 1 and 2, as well as relevant literature (Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010) arrows failed to capture participants attention compared to the other two social cues. As was discussed extensively in the previous experiments (see discussions), a potential reason for the low dwell times on the arrows, are due to the importance social cues pose for the observer (e.g., understanding actors' intentions). Another explanation is that arrows are simple (possess fewer characteristics) and thus easier to process. Therefore, there is no need for the observer to spend a lot of their time looking at them. Potentially, the nature of the task (e.g., looking for known target) can be the cause for the shorter dwell times on the arrows. All these explanations, however, should be considered with caution, because of the issue with data recording on some trials.

One major difference between the three experiments was the cueing effect between social and symbolic cues, measured by the saccades to the target. In the previous literature (e.g., Birmingham et al., 2009; Hermens & Walker, 2015) and experiments of this chapter, eye movements' patterns have been explored under a free viewing image or by asking the observers to locate the person presented in the scene. However, it was still unclear if social and symbolic cues embedded in natural scenes, have a strong cueing effect when a visual search for a specific item is performed. In this experiment, results from the task suggest that both gaze and pointing cues led more successfully to the targeted item compared to the arrow. These findings are in agreement with Burton et al. (2009) where both peripherally presented pointing hands and rotated heads produced the same effects on observers' attention.

However, saccades findings contradict those from Experiment 1 and 2 where arrows showed a higher proportion of saccades to the target. A logical explanation can be the difference in the tasks. In the first two experiments eye movements were not under the influence of any specific instructions (e.g., find a target). This could have allowed arrows to show a stronger shifting of attention towards the target (see discussions in Experiment 1 and 2 for an explanation). However, in this study cueing effects from arrows was fairly weak. There is a wide range of studies which have suggested that centrally presented gazing cues produce a strong cueing effect when compared with another social cue or an arrow (Driver et al., 1999; Friesen & Kingstone, 1998; Langton & Bruce, 1999). When presented in the periphery, head, pointing and arrow cues showed strong cueing effects, due to their distinct shape, compared to a set of eyes (e.g., Hermens, 2017) The current findings on the cueing effect towards the target, comes into an agreement with the studies presenting cues at and away from fixation. Therefore, the solution should be looked somewhere else.

A potential explanation can be given in relation to the decreased numbers of the percentages of the cueing effects. Going back to the saccades for experiment 1 and 2, percentages of the saccades to the target reached to more than 50%. However, in this experiment, percentages of the saccades leading to the target were substantially lower. Therefore, these reductions might be another indication that cues were ignored and not used. One question still remains why social cues elicited more attentional shifts compared to arrows. In the literature it has been argued that attention shifts by symbolic cues, depend on a goal driven (voluntary) attentional mechanism as participants realise that using the cue is not useful to locate the target (Jonides, 1981; Tipples, 2008). Social cues do not seem not to fall into the goal driven category (maybe due to their biological relevance). Therefore, it can be assumed that an exogenous (involuntary) and stimulus driven mechanism is taking place (Friesen & Kingstone, 1998; Posner, 1980).

Experiment 1 and 2 also explored the cueing effect towards the environment. In both experiments social and symbolic cues led more frequently towards the other elements of the environment and less to the targeted object we have assigned. The possible reasons which led to these effects were extensively discussed in Experiment 1and 2 discussion. However, in this experiment, cueing effect towards the environment showed different results, with cues directing participants' gaze more effectively to the target. So, what made the cueing effect decrease towards the environment? An obvious reason is that the previous two studies were free viewing where participants' eye movement patterns were not influenced towards certain elements (target or cues) in the scene. Whereas in this study, participants had to locate a specific target in the scene in a restricted amount of time. These restrictions could have biased any intention from observers to explore the other elements in the scene (apart from cues and the target).

Although fixations on the background of the scene were not as frequent as in Experiments 1 and 2, they were still fairly frequent. If results from reaction times, dwell times and saccades are taken into account, we can see that while searching for a target, cues were not consistently fixated. Moreover, shifting of attention did not lead exclusively to the target but also to the background. These patterns seem to be in agreement with studies that presented cues in more cluttered environments (e.g., more items and cues in the display) rather than a single cue or a target embedded in a blank display (Hermens, 2017).

2.5 Conclusion

This chapter investigated the effects of social and symbolic cues on observers' attention when cues are embedded in static natural scenes. Three experiments were conducted: (1) single cues and free viewing, (2) single and double cues and free viewing, (3) single cues and visual search. In most aspects, the results from the three experiments were consistent. Gaze cues consistently captured participants' attentions even when more than one cue was present in the scene. Only when paired with a pointing cue, gaze cues had less of an effect on observers' eye movements. Results, however, depended on how the regions of the cues were defined (just the head, or the arm or the entire body) and future studies should be clear on how cue regions are analysed. The effects of the arrows on saccades away from the cue depended on the task and were stronger during free viewing. This chapter has explored static scenes, which are already a step forward from traditional cueing paradigms, by placing cues in a natural scene. The next chapter will move one step further by examining cueing by dynamic scenes (videos).

Chapter 3

Exploring the effects of dynamic social and symbolic cues in natural scenes

3.1 Introduction

Studies of social attention have suggested that gaze cues strongly influence people's attention in the gazed-at direction, even when they are instructed to look in the other direction (e.g., Bindemann, Burton, Hooge, Jenkins, & de Haan, 2005; Theeuwes & der Stigchel, 2006). The effects of gaze cues are suggested to be strong even when the gazed-at direction is from a peripherally presented target, with observers showing a strong bias towards eyes' direction (e.g., Bayliss & Tipper, 2006b; Friesen et al., 2005, 2004; Sato, Okada, & Toichi, 2007; Tipples, 2008). Similar effects have been found for another social cue (pointing) and symbolic cue, such as an arrow (e.g., Hommel et al., 2001; Langton & Bruce, 2000; Ristic et al., 2002; Tipples, 2002) suggesting that the cueing effect is not unique to gaze cues.

Studies so far have presented the cues in isolation (i.e., in an otherwise empty display) and frequently at fixation. More recent studies (see also Chapter 2 of this thesis) have proposed to explore the effects of social and symbolic cues when they are part of natural scene environment to gain a better understanding of social attention in the real-world (e.g., Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010). While previous studies as well as the results presented in Chapter 2 provide valuable information about the influence of social and symbolic on observers' attention, these studies all used static cues. In day – to-day encounters, such static cues are uncommon, particularly for social cues (e.g., one rather observes a head turning rather than an already turned head).

The literature of social attention seems to have avoided videos as stimuli. For example in several studies presented eyes are covered before the gaze cue is presented to observers or averted gazing eyes are presented immediately after the fixation point (Bayliss et al., 2005; Hermens & Walker, 2010). This avoidance of videos as stimuli may be due to the possible confound between the effects of motion cues and the social aspect of the cues (Farroni et al., 2000). Evidence from neuroimaging studies have supported this idea by showing that the same neurons are sensitive to both biological motion (Oram & Perrett, 1994) and social cues (Perrett et al., 1992, 1985).

Few studies have attempted to work around the confound between motion cues and social attention effects in an attempt to bring the work of social attention closer to realistic situations (Bayliss et al., 2005; Farroni et al., 2000; Hermens & Walker, 2012; Kuhn & Tipples, 2010; Ricciardelli et al., 2002; Rutherford & Krysko, 2008; Swettenham, Condie, Campbell, Milne, & Coleman, 2003). A typical strategy is to compare the effects of stationary eyes in a head moving in the opposite direction and shifts of gaze in a stationary head (Bayliss et al., 2005; Hermens, 2015). Findings from these studies are mixed. In typical and atypical (ASD) populations, attention shifts seem to be influenced by the direction of the gaze rather than general motion direction (i.e., the direction of motion of the head; Bayliss et al., 2005; Hermens, 2015; Rutherford & Krysko, 2008). In contrast, young infants (16 to 21 weeks old) show an attentional shift in the direction of motion rather than gaze (Farroni et al., 2000). Dynamic cues seem to have similar effects as static cues, because studies measuring manual response times using static eyes (initially covered and later revealed) showed similar effects (Bayliss et al., 2005). More recent studies (e.g., Foulsham et al., 2010) have managed to explore how people's gaze allocation is affected in the presence of a dynamic gaze cues by using a more close to reality scenario: a dynamic social situation (for more details please see section 1.3. in general introduction). Results from their study revealed that attention on the eyes and faces was stronger compared to the other elements in the video (e.g., background, people's torso).

While some studies have attempted to study the effects of dynamic gaze cues, no studies have tried the same for pointing and arrow cues. Some evidence for similar effects of such cues is provided by neuroimaging studies, which have suggested that the same region (STS region) is activated when people viewing hand gestures (Grèzes, Costes, & Decety, 1999; Nakamura et al., 2004) and hand movements (Bonda, Petrides, Ostry, & Evans, 1996;

Pelphrey, Morris, & McCarthy, 2004; J. C. Thompson, Hardee, Panayiotou, Crewther, & Puce, 2007). This area is also active when directional movement of the eyes and head is observed (T. Allison et al., 2000). There is no behavioural studies so far investigating the effects of a dynamic pointing hand can have on the observers' attention and how motion can influence these effects. However, there is a plethora of studies (e.g., Liszkowski, 2005; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004) exploring dynamic pointing hands in children. For instance, a study by Rohlfing et al. (2012) have suggested that pointing dynamic hand is capable to orient children's attention from the age of 4.5 months old. The same study suggested that motion is a necessary element but not sufficient to direct children's attention and at early age humans can understand the directional information pointing gestures communicate.

Several studies have suggested that eye movements strongly depend on the task of the observer (e.g., Fletcher-Watson et al., 2008; Itier et al., 2007; Yarbus, 1967). For example, different eye movement patterns were found when judging the age or the gender of the people in the scene. A more recent study (Hermens & Walker, 2015) specifically looked at the effects of gaze and pointing cues under different tasks (free viewing versus a memory task) that kept presentation times of the images constant and found that dwell times on the cues was similar for both tasks. Although there is a plethora of evidence for the influence of task in static images, to our knowledge there is no existing literature exploring task effects on dynamic cues. Therefore, it is unclear if gaze behaviour on the cues differences between free viewing and task will persist.

To better understand the effects of dynamic cues, the present study examines the effects of social (head-turning and pointing) and symbolic cues (an arrow appearing) on eye movements of observers. In contrast to (above study), the present study will examine the effects of these cues when they are embedded in a natural scene to approach day-to-day viewing as closely as possible. Considering the above evidence, this study has the following aims. Previous studies using static cues have demonstrated the strength that social cues, and particularly gaze cues, have to bias people's attention towards them (e.g., Birmingham et al., 2009; Fletcher-Watson et al., 2008). This effect is persistent even when compared with another social cue, such as pointing or a cue with no biological relevance (e.g., arrows). However, for dynamic social and symbolic cues it is still unclear if this effect will persist. For this reason, the first aim of this study is to compare and explore dynamic social and symbolic cues (embedded in natural scenes) in order to understand which cue will prevail over capturing people's attention.

Second, while previous studies (e.g., Bayliss et al., 2005) compared static and dynamic cues suggesting to an extent similar cueing effects for the gaze cues, this effect might be a result from the way cues were presented (isolation and at fixation). Therefore the second aim is to compare cueing effect between dynamic arrow signs and social cues when cues are presented in a natural scene and away from fixation. In contrast to previous studies such as Bayliss et al. (2005), this experiment will present both head and eyes following the same direction for both social cues.

The final aim is to compare two different tasks to explore any differences in attending the three dynamic cues. To achieve this a free viewing task will be compared with a memory task, where participants will be asked to memorize the scene for a follow up memory test, similar to that used by Hermens and Walker (2015). In Chapter 2 (see Introduction) it was argued that free viewing task at one hand is ideal to explore the effects of social and symbolic cues on observers' attention, but on the other hand has the possibility for the participants to find no purpose of just observing a scene. In Experiment 3 (Chapter 2) a more active and closer to real life scenario was used (visual search task). In the current experiment, a different task will be used (memory task) and compared with a free viewing task. The memory task was chosen as a compromise between an active task (such as visual search), and a task where the presentation time of the image could be matched with that during free viewing. Visual search namely ends when participants have found the target and they may switch to free viewing when asked to look at the image longer. As in a visual search task, during a memory task participants acquire a goal to memorize the scene as their memory will be assessed afterwards. Under these instructions (e.g., memorize the content of a scene) participants' eye movements are also manipulated. Participants have to visit and memorize as many points in the scene as possible, for them to retain these points in their memory. Therefore, as in a visual search task, a memory task can also show how eye movements towards the cues differ between a without purpose task (free viewing) and a more active task (memory task). Based on past results (e.g., Hermens & Walker, 2015), task differences might be expected only for the saccades leaving the cue, while dwell time on the cues are expected to be similar across the two tasks.

To address these aims, dynamic social cues (pointing gesture, a head with same direction and eyes turning towards the direction of a targeted object) that are embedded within short videos of natural scenes, will be used. Social cues will be provided by real people with their body in full view. Arrows will be presented at the same location as the actors and be distinct in presentation to give equal strength to the two social cues. A challenging element of this study was to create a dynamic version of an arrow cue. In most situations arrows are stationary and do not move. One strategy could be to have part of the arrow presented in the scene as in the study of Crostella, Carducci, and Aglioti (2009) (part of the arrow head was missing). However, this might have created a bias towards this element in the scene. Therefore, in this study, an arrow will appear at the same time actors' initiated their motion and disappear when actors started going back to their initial position (also see Methods section). Based on previous studies (e.g., Crostella et al., 2009; Hermens & Walker, 2012) it is expected that actors' faces will strongly capture observers' attention whereas arrows will lead to a weaker looking time. Cueing effect has being suggested to be strong to both social cues (Crostella et al., 2009) therefore we expect a strong competition between gazing and pointing.

3.2 Experiment 4

3.2.1 Methods

3.2.1.1 Participants

Thirty two (6 male and 26 females; 18 to 43 years old, mean = 23.97, SD= 6.25) undergraduate students and members of staff from the University of Lincoln took part in this study. All had normal or corrected-to-normal (contact lenses) vision. All participants provided written consent prior the beginning of the study that was approved from the University of Lincoln ethics committee.

3.2.1.2 Apparatus

Participants were seated in front of a ViewSonic VX2268WM flat screen (1280 × 1024 pixels resolution and 60Hz refresh rate) at a distance of 80 cm (approximately 26° x 20° of visual angle), restricted by a chin rest. Eye movements were recorded with an Eyelink 1000 (SR Research Osgood, ON, Canada) desk mounted eye tracker. Eye tracking was performed monocularly using the right eye of the participants. Stimuli were displayed using the Experimental Builder software (SR Research Osgood, ON, Canada), while a second PC recorded subjects' eye movements at a sampling rate of 1000 Hz. All participants' eye movements were tracked using the combined corneal and pupil reflection modes of Eyelink system. The videos used in this experiment were recorded with a Hero 5 Session GoPro camera (San Mateo, California, U.S.).

3.2.1.3 Stimuli

Forty-four videos were used, showing 11 different indoor scenes from the University of Lincoln (Sarah Swift building). Four type of videos were created for each scene (Figure 3.1), showing the scene with (1) a gaze cue; (2) a pointing cue; (3) an arrow cue and (4) without any cue. Gaze cue videos clips showed an actor in four stages: (1) standing, for 1 second, (2) initiating a movement (turning head and body towards a targeted item) lasting approximately 3 to 3 ½ seconds (depend on actors' different movement) , (3) standing for 3 seconds in this position, and (4) turning back to the starting position. The pointing cue video clips showed actors' movements with the same four stages, but also had a pointing movement together with turning the head and body. Actors moved in such a way that they always simultaneously turned their head and eyes. The third type of videos contained an arrow cue. To better mimic the gaze and pointing cue videos, the arrow was introduced after 2 seconds into the video (it appeared at this time) and disappeared after 4 seconds . To avoid gaze and pointing cues to strongly depend on the actor, five different actors (including the experimenter) were used, all having different body dimensions and different movement speeds. As a consequence, the duration of the videos varied between 8 to 8 1/2 seconds.

The target object differed across scenes, but was consistent across cues for that scene. Videos were scaled to a size of 1280 x 720 (26° x 14° of visual angle) and were edited with Adobe Premiere Pro. For presentation within the eye tracker, videos were converted to the xvid format and mp3 mode. The same video editing software, was used to embed the arrow in the scene. Participants either performed free viewing or a memory test. For the memory test, cued objects from each scene were presented among three similar items on a power point presentation in a separate pc.

3.2.1.4 Design

Half of participants performed a memory task, whereas the other half performed a free viewing task. This made the task a between-subject factor. All participants were exposed to the same videos and number of cues. As a consequence, participants saw the same scene four times (once for each cue). To avoid influences of this repetition on the average data and avoid the subsequent presentation of the same cue on successive trials, the order of the cues within the videos was systematically varied across participants by creating four blocks (unknown to the participants). Each block included an equal number of no-cue, gaze, pointing and arrow cue videos, and each scene appeared only once in each block. The order of the videos within each block was randomized for each participant. This design allows two analyses: (1) an analysis of all data, making cues a within-subjects factor (with possible repeated exposure effects, but not in the average data), and (2) an analysis of the first block of trials, making cues a between subjects factor (excluding the possible repeated exposure effects).

3.2.1.5 Procedure

The procedure for tasks was highly similar. Before taking part, participants received written and verbal instructions about the experiment. Subjects performing the memory task were instructed to pay attention to the scene, because a memory test about the videos would follow at the end of the experiment. Participants performing the free viewing task were asked to simply look at the videos. To minimize head movement, participants were instructed to place their head on a chin rest. The eye tracker was calibrated with a nine-point calibration. On most occasions, calibration was successful in a first or second try, but if needed, calibration was repeated until the recorded fixation positions were aligned with the same three-by-three grid on which the fixation targets were present. Each trial started with a drift correction target presented above or below the position of the video (at the vertical midline), ensuring not to bias participants' gaze direction towards a particular region of the video (e.g., the center). Drift correction also ensured realignment of the recorded eye position if needed. Each video was presented for approximately eight seconds. After 22 trials participants were offered a break and where needed a recalibration was performed. At the end of the experiment, participants performing the memory task were given 11 slides to test their memory of the scenes. These data were not analysed, as this task only served to reinforce the task instruction.

Participants were then debriefed about the purpose of the experiment and were offered the opportunity to ask questions.

3.2.1.6 Data Analysis

As before, the raw eye movement data were automatically parsed into fixations and saccades and blinks using the Eyelink's parser software (SR Research). The data viewer software (SR Research) was then used to assign each fixation to the appropriate region of interest (ROI). ROIs were defined with this software and included actors' body (trunk and legs), head and hand (for pointing cues) and arrows' sign for the arrow cues. The arrow ROI was defined as two regions, depending on the temporal position of the video frame in the video. One region was aligned with the time actors were turning (turning phase, Figure 3.1). The second region matched the period in which the actors were fully turned towards the target item (cueing phase, Figure 3.1). As these two stages were chosen to analyse these data, no cue condition was not used in the analysis. Regions were fitted as tightly as possible to the cues to achieve a more precise estimation to where participants' fixations fell. One challenge for analysing fixations in dynamic cues, was when fixations overlapped between the turning and cueing periods. This usually occurred when motion was about to stop and actors were assuming their cueing position. As data viewer software, cannot compensate for overlapping fixations, manual coding was required (similar to the manual coding used in Chapter 4). Experimenter parsed all videos and assigned the under question fixations to the appropriate phase. For instance if 80% of one fixation fell into the turning phase and 20% fell into the cueing phase, fixation was assigned to the turning phase. Fixations that fell 50% to the turning phase and 50% to the cueing phase were excluded.

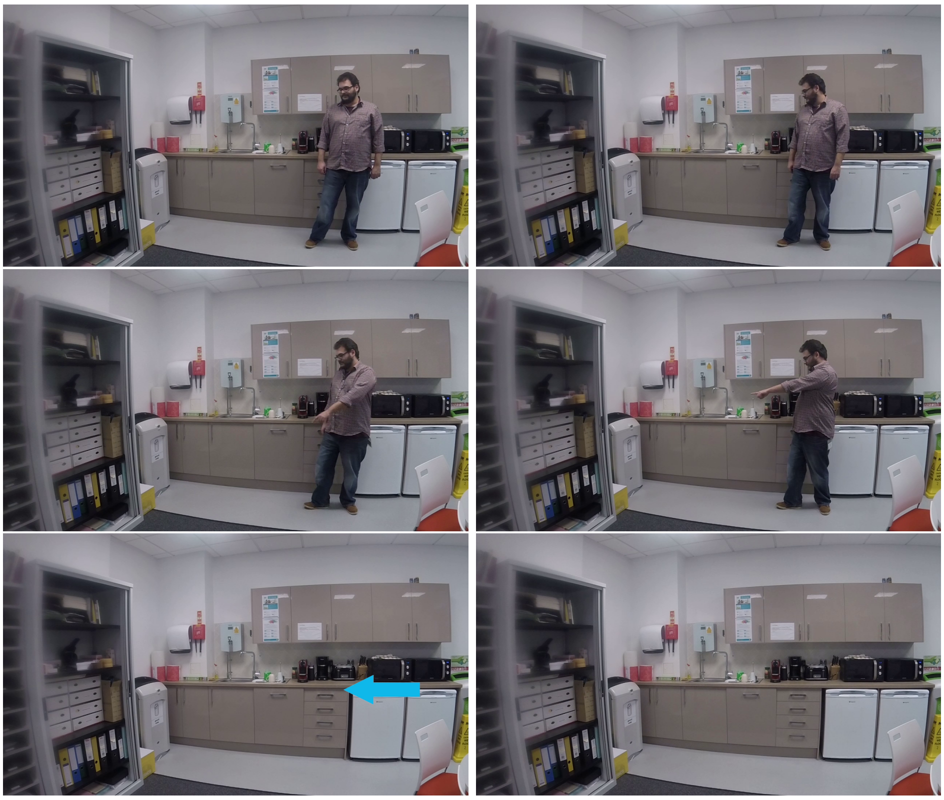

Fig. 3.1 Screenshots of stimuli and the phases (turning and cueing) used in Experiment 4. Left image (top) showing a person start gazing (gazing condition/turning phase). Right image (top) shows the same person gazing at the target (gazing condition/cueing phase). Left image (middle) showing a person start pointing (pointing condition/turning phase). Right image (middle) shows the same person pointing at the target (pointing condition/ cueing phase). Left image (bottom) shows an arrow pointing at the target. Right image (bottom) shows a "No Cue" scene.

Because the main interest was in the effect of the presence or absence of movement, only the stages where the actors were initiating and stopping motion were analysed. For statistical comparisons between different conditions, linear mixed-effects models were used (lme4 package by Bates et al., 2014) and Bonferroni correction was applied for multiple comparisons. For the linear mixed-effects models only the χ^2 and p-value will be reported. Similar to the experiments in Chapter 2, two analyses will be reported: (1) a initial analysis where all the trials were used and (2) a secondary analysis (only for the dwell times on the cues and saccades) where only the first block of trials was used.

3.2.2 Results

3.2.2.1 Dwell Times on the cues

Figure 3.2 shows the average dwell times on the cue for the two different tasks (memory, free viewing) and the two parts of the movement of the actor (turning and cueing), pooled across the scenes. Please note that dwell times on the cues are expressed as a percentage of the total trial time. The plot suggests that while the actor is moving or the arrow appearing ('Turning phase' in the plot), dwell times on arrow cues were slightly longer than those on pointing hands and clearly longer than those on gaze cues. During the cueing stage (the period in which the actor is gazing or pointing at the target without movement; or when the arrow is present) dwell times were longest on the gaze cue, followed by pointing cues and arrows. Except for dwell times on the gaze cue during the 'turning' phase, these dwell times do not seem to depend on the task (Free view versus Memory in the plot).

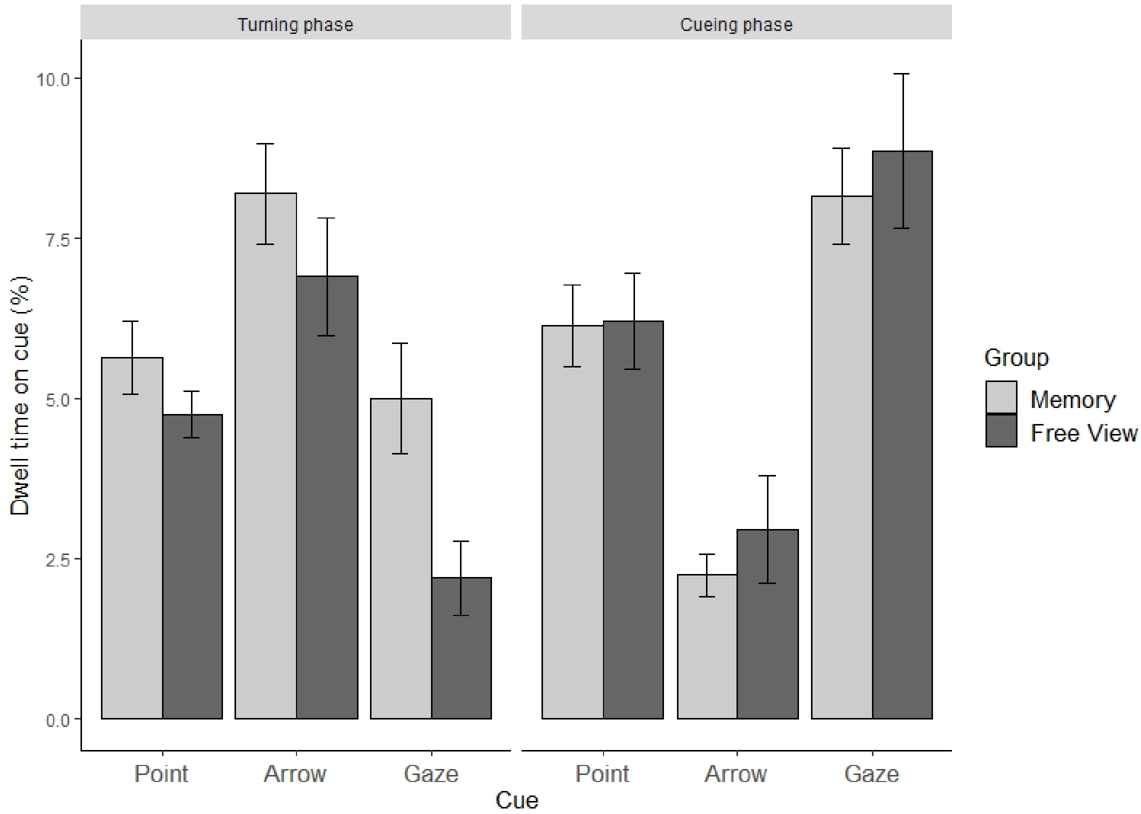

Fig. 3.2 Dwell Times on the cue for the two tasks (Memory, Free viewing) and the two motion stages (Turning phase= during turning and Cueing phase = during cueing), as a percentage of the total duration of the interval. Averages were obtained by first pooling across scenes and then across participants. Error bars show the standard error across participants.

These observations are confirmed by a linear mixed-effects analysis. A significant interaction was found between cue type, task and motion stage (participants and scenes as random factors) ($\chi^2(6) = 295.53, p < 0.0001$), suggesting that the pattern of cue dwell times differs significantly between tasks and motion stages. First block of trials analysis confirmed these results ($\chi^2(6) = 85.18, p < 0.0001$). A significant two way interaction between the two tasks and the two motion stages for each cueing condition was computed, showing significant results for the arrow cues ($\chi^2(1) = 5.17, p = 0.02$) and gaze cues ($\chi^2(1) = 13.20, p = 0.0003$) but not for pointing cues ($\chi^2(1) = 1.58, p = 0.21$). Similar results were found from the first block of trials analyses (gaze: $\chi^2(1) = 2.97, p = 0.04$; arrow: $\chi^2(1) = 0.51, p = 0.02$ and pointing: $\chi^2(1) = 1.99, p = 0.16$).

In subsequent analyses, the effects of task and cue within each motion stage were investigated.

Analysis on the turning motion stage

Linear mixed-effects model analysis (using participants and scenes as random factors) showed no significant interaction between cue and task during the turning phase ($\chi^2(2) = 5.05, p = 0.08$). When first block of trials was analysed similar results were found for the interaction between cue and task during the turning phase ($\chi^2(2) = 2.11, p = 0.11$). Linear mixed-effects analysis for the main effect of the task on cueing condition, showed a significant difference for the gaze cues ($\chi^2(1) = 7.90, p = 0.01$) but no significant for the arrows and pointing (arrows: $\chi^2(1) = 1.19, p = 0.27$; pointing: $\chi^2(1) = 3.01, p = 0.08$). First block of trials analysis showed similar results. These results suggest that the only cue that was affected by the task during the turning motion stage was the gazing cues. A possible explanation can be that during this motion stage, gazing cues speed was brief (compared to pointing), allowing for the task to have an effect on participants' eye movements.

Analyses on the memory task during the turning stage

Differences between the three cues for the memory group (turning stage) were also computed. Linear mixed-effects analysis of all the trials as well as first block of trials analysis showed significant difference (All trials analysis: $\chi^2(2) = 14.22, p = 0.0008$; First block of trials analysis: $\chi^2(2) = 2.83 = 0.006$). Comparison mixed effect analyses of the cues for the memory group in turning stage showed that longer dwell times for arrows (arrow vs gaze: $\chi^2(1) = 12.00, p = 0.001$; arrow vs point: $\chi^2(1) = 6.95, p = 0.01$) and no significant differences between gaze and point ($\chi^2(1) = 1.22, p = 0.27$). When only the first block of trials was analysed similar results were found between the arrow and pointing cues

($\chi^2(1) = 3.01, p = 0.02$) and between gaze and pointing cues ($\chi^2(1) = 1.50, p = 0.22$), but not between arrow and gaze cues ($\chi^2(1) = 1.82, p = 0.18$).

Analyses on the free viewing task during the turning stage

Similar analyses were carried out for the free viewing task (turning stage). Linear mixed-effects model analysis on the differences in dwell times on the cues for the three cueing conditions showed significant differences ($\chi^2(2) = 56.50, p < 0.0001$). These difference were also found when only the first block of trials was analysed ($\chi^2(2) = 30.51, p = 0.004$). Mixed effect model comparisons of the cues for the free viewing task showed longer dwell times for the arrows compared to the two social cues (arrow vs gaze: $\chi^2(1) = 50.74, p < 0.0001$; arrow vs point: $\chi^2(1) = 8.47, p = 0.004$) and longer dwell times for the pointing compared to gazing ($\chi^2(1) = 28.08, p < 0.0001$). First block of trials analyses showed similar results.

Analyses on the cueing motion stage

Final analyses for dwell times on the cues explored the interaction between the cues and the two task groups for the cueing motion stage, showing a significant interaction ($\chi^2(5) = 264.76, p < 0.0001$). When only the first block of trials was analysed, similar significant results were found ($\chi^2(5) = 71.11, p = 0.003$). Linear mixed-effects model analysis for the effect of the task for the three cueing conditions showed no significant differences (arrows: $\chi^2(1) = 0.64, p = 0.42$; gaze cues: $\chi^2(1) = 0.01, p = 0.95$; point cues: $\chi^2(1) = 0.002, p = 0.96$). Same results were found when only the first block of trials was analysed (gaze cues: $\chi^2(1) = 0.20, p = 0.70$; arrows: $\chi^2(1) = 0.44, p = 0.33$ and pointing cues: $\chi^2(1) = 0.07, p = 0.81$). These results contradict the findings for task effects on the gaze cues dwell time during the turning motion stage. It can be speculated that this difference might be an after effect of the previous motion stage (Turning stage). When gaze cues stopped moving, participants in both task groups showed equal interest to actors in the scene. Therefore what was observed, in the previous stage for the gazing might be a combination of motion and task effect.

Analyses on the memory task during the cueing stage

Exploring the differences on dwell times between the three cues for the memory task showed significant differences ($\chi^2(1) = 168.22, p < 0.0001$). These results were confirmed from the analysis of the first block of trials ($\chi^2(1) = 73.59, p = 0.001$). Mixed effect comparison analyses exploring the differences between the dwell times on the cues for the three cueing conditions under the memory task showed longer dwell times for the gazing cues compared to arrows ($\chi^2(1) = 142.97, p < 0.0001$) and to pointing ($\chi^2(1) = 40.62, p < 0.0001$) and longer dwell times for pointing compared to arrows ($\chi^2(1) = 67.98, p < 0.0001$). First block of trials showed similar results (gaze vs arrows: $\chi^2(1) = 60.28, p = 0.002$; gaze vs pointing: $\chi^2(1) = 25.67, p = 0.01$ and pointing vs arrow: $\chi^2(1) = 35.94, p = 0.03$).

Analyses on the free viewing task during the cueing stage

Similar analyses were computed for the free viewing task showing significant differences between dwell times on the cues for the three cueing conditions ($\chi^2(1) = 104.13, p < 0.0001$). Comparing the three cues with Linear mixed-effects model analyses result suggest same pattern with the memory task group, with gazing cues showing longer dwell times compared to pointing ($\chi^2(1) = 31.52, p < 0.0001$) and arrows ($\chi^2(1) = 87.61, p < 0.0001$) and pointing showing longer dwell times than arrows ($\chi^2(1) = 32.57, p < 0.0001$). Similar results were found when only the first block of trials was analysed.

3.2.2.2 Trials with fixations on the cue

Dwell times on cues were relatively short, and one may wonder how much these short dwell times reflect participants ignoring the cue (resulting in zero dwell times on those trials). For this reason, Hermens and Walker (2015) also looked at how many trials had at least one fixation on the cue, which will also be used here.

Figure 3.3 shows that the percentage of trials with at least one fixation on the cue while the actor was turning suggesting that task and motion stage had no clear effect on how often the cue was fixated at least once.

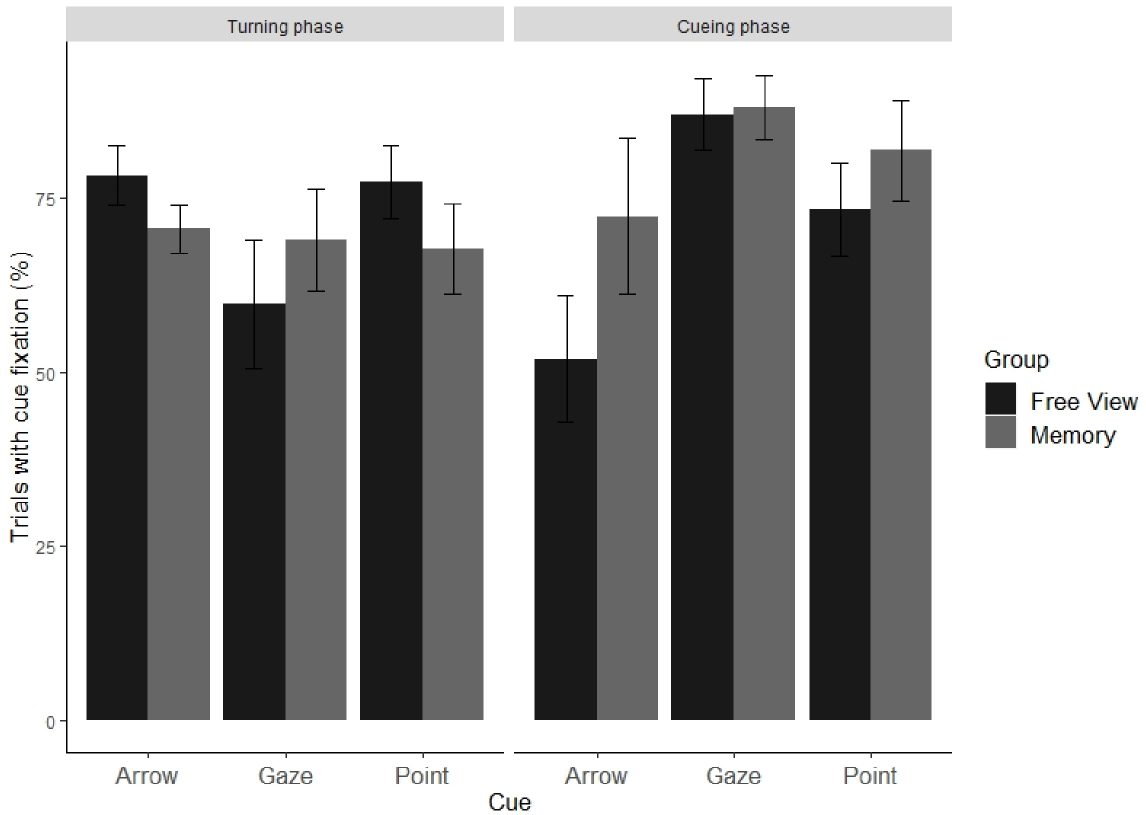

Fig. 3.3 Percentage of trials with at least one fixation on the cue for the two task groups (free viewing, memory) and motion stages (Turning phase= during turning, Cueing phase = during cueing). Error bars represent the standard error of the mean across participants.

Interaction analysis

Linear mixed-effects model analysis for the interaction between the three cueing conditions, two motion stages and groups, showed no significant differences ($\chi^2(7) = 10.01, p = 0.19$).

Within stage analysis

Linear mixed-effects model analysis on the percentages of trials with at least one fixations on the cues, using participants and scenes as random factors, for the turning motion stage, showed no significant difference for the cues between the groups (arrow: $\chi^2(1) = 0.13, p = 0.72$; pointing: $\chi^2(1) = 0.006, p = 0.94$ and gazing: $\chi^2(1) = 2.11, p = 0.15$). Detailed analysis in each group for the same motion stage showed no difference between the cues for the free viewing group (gaze vs point: $\chi^2(1) = 1.10, p = 0.30$; gaze and arrow: $\chi^2(1) = 0.77, p = 0.38$; point vs arrow ($\chi^2(1) = 0.16, p = 0.69$) and no difference between the cues for the memory group (arrow vs point: $\chi^2(1) = 0.45, p = 0.50$; gaze vs arrow: $\chi^2(1) = 1.03, p = 0.31$; gaze vs point: $\chi^2(1) = 0.12, p = 0.73$).

There are substantial numbers of trials without a fixation on the cue (around 25% for the pointing and arrow cues, and around 40% for the gazing). Figure 3.5 also shows that the percentage of trials with a fixation on that cue while cueing towards the targeted item is higher for gazing cues than pointing cues(mixed effect logistic regression: $\chi^2(1) = 9.42, p < 0.002$) and lower for arrows than gazing (mixed effect logistic regression: $\chi^2(1) = 11.80, p = 0.001$). There are substantial number of trials without a fixation on the cue (25% for pointing and arrows and around 35% to 40% for gaze) for both groups.

Exploring the differences on the percentages of trials with a fixation on the cue between the two groups and cues for the cueing motion stages analysis (linear mixed-effects models) showed similar results with cues showing no significant differences between the two groups.

Mixed effect comparison analysis for the cues in each group showed similar patterns with the turning motion stage only for the two social cues. Arrows showed significant differences on the percentages of trials on the cues, when compared with other two social cues for the memory group (arrow vs gaze: $\chi^2(1) = 7.50, p = 0.006$; arrow vs point: $\chi^2(1) = 1.99, p = 0.05$) as well as the free viewing group (arrow vs gaze: $\chi^2(1) = 6.77, p = 0.01$; arrow vs point: $\chi^2(1) = 7.38, p = 0.007$). Again there is a substantial number of trials without a fixation on the cue but it seems to be higher (at least for the gaze cues) compared to the turning motion stage (approximately 25 to 50% for arrows, 10% for gaze and 25% to 20% for pointing

3.2.2.3 Dwell Times on the Head, Body and Arm

In the analysis so far, the gaze and pointing cues were assumed to include the head and arm, as well as actors' entire body. To examine which region participants fixate for two types of cues, Figure 3.4 plots the dwell times for these separate regions and for the two motion stages and the two groups. Similar to the experiments in Chapter 2, this plot shows the percentage of total time spent on the three sub-regions, given that participants looked at the different body parts (at least one fixation to the region). The graph suggests that while actors were turning, participants fixated the head most for both types of cues followed by the body for the gazing cues and finally the arm for the pointing cues. In addition, when social cues assumed their cueing position the plot suggests that participants fixated the head most for the gazing cues and less for the pointing cues; the hand for the pointing cues and finally the body.

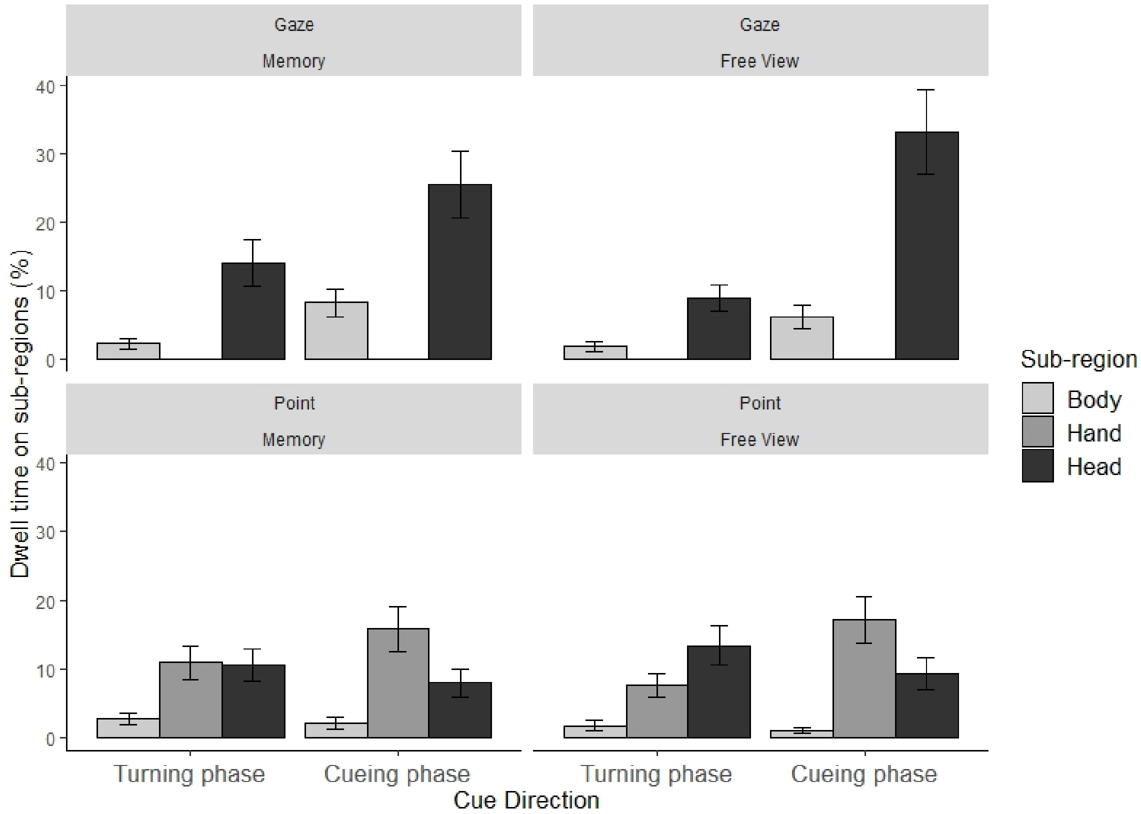

Fig. 3.4 Dwell time on the three sub-regions for the two social cues, two cue motion stages (Turning phase = during turning and Cueing phase = during cueing) and task groups (memory and free viewing). Error bars represent the standard error of the mean across participants.

Interaction analysis

A linear mixed-effects model analysis (allowing for missing rows due to participants sometimes not fixating the people in the images) confirms that there was an interaction between the two social cues; the two motion stages; two tasks and the two sub-regions ($\chi^2(5) = 48.87, p < 0.0001$). A subsequent three way interaction analysis between the two motion stages, task groups and sub-regions (hand only included in pointing cues) per cue type showed significant differences for both gaze ($\chi^2(4) = 54.92, p < 0.0001$) and pointing cues($\chi^2(7) = 62.67, p < 0.0001$). Two way interactions were conducted between the two sub-regions (head and body) and the two groups per motion stage, showing a significant results for the cueing stage ($\chi^2(3) = 30.98, p = 0.05$) and turning stage ($\chi^2(3) = 184.82, p < 0.0001$).

Within task analysis

As the cues were explored under different motion stages (turning and cueing) it is interesting to explore the difference the two body regions had in each stage and for each task group. While turning distribution of fixations across the ROIs and for each task was different for gazing and pointing cues (Free viewing: $\chi^2(1) = 37.99, p < 0.0001$; Memory: $\chi^2(1) = 2.92, p = 0.001$). The main effect of the two sub-regions (three for pointing) for each social cue and group showed significant differences for gazing (Free viewing: $\chi^2(1) = 33.57, p < 0.0001$; Memory: $\chi^2(1) = 68.78, p < 0.0001$) and pointing (Free viewing: $\chi^2(1) = 74.51, p < 0.0001$; Memory: $\chi^2(1) = 45.70, p < 0.0001$) in both groups. Pairwise linear mixed-effects model comparisons, examining the dwell times on the regions for both social cues and for the two groups are shown in Table (3.1).

A same set of analyses were carried out for the cueing motion stage. While cues assumed their cueing position distribution of fixations across the ROIs and for each task was different for gazing and pointing (Free viewing: $\chi^2(2) = 170.79, p < 0.0001$; Memory: $\chi^2(2) = 264.47, p < 0.0001$). The main effect of the two sub-regions (three for pointing) for each social cue and group showed significant differences for gazing (Free viewing: $\chi^2(1) = 162.92, p < 0.0001$; Memory: $\chi^2(1) = 97.49, p < 0.0001$) and pointing (Free viewing: $\chi^2(1) = 114.99, p < 0.0001$; Memory: $\chi^2(1) = 95.36, p < 0.0001$) in both groups. Pairwise linear mixed-effects model comparisons examining the dwell times on the regions for both social cues and for the two groups are shown in Table (3.1).

Table 3.1 Results from paired comparison (χ^2-value, p-value) for the dwell times on the three sub-regions in each cueing condition and for each task group.

Cue	Comparison	Group	χ^2-value	p-value
Gaze (turning)	Head/Body	Memory	68.76	<0.0001
Point (turning)	Head/Body		42.51	<0.0001
	Head/Hand		0.002	0.97
	Hand/Body		38.72	<0.0001
Gaze (cueing)	Head/Body		97.40	<0.0001
Point (cueing)	Head/Body		25.18	<0.0001
	Head/Hand		28.36	<0.0001
	Hand/Body		90.88	<0.0001
Gaze (turning)	Head/Body	Free View	35.57	<0.0001
Point (turning)	Head/Body		70.01	<0.0001
	Head/Hand		15.02	0.001
	Hand/Body		32.67	<0.0001
Gaze (cueing)	Head/Body		162.92	<0.0001
Point (cueing)	Head/Body		51.65	<0.0001
	Head/Hand		21.99	<0.0001
	Hand/Body		115.6	<0.0001

3.2.2.4 Dwell Times on the target

Because of the possible confound between ROI size, ROI location and dwell times, which is close to impossible to control for in natural scenes, dwell times on cues are difficult to interpret. In contrast, dwell times on cued objects are not expected to be influenced by such factors, as the cued object is always the same across cueing conditions. Figure 3.5 shows these dwell times on the cued object for the three different conditions, task groups and the two motion stages (turning and cueing), pooled across the scenes. The plot suggests substantially longer dwell times on the target for pointing cues during the cueing stage (‘Cueing phase’), but hardly any differences between cues during the turning stage (‘Turning phase’). Similar to the dwell times on the cues analysis, dwell times on the target are also expressed as a percentage of the total trial time.

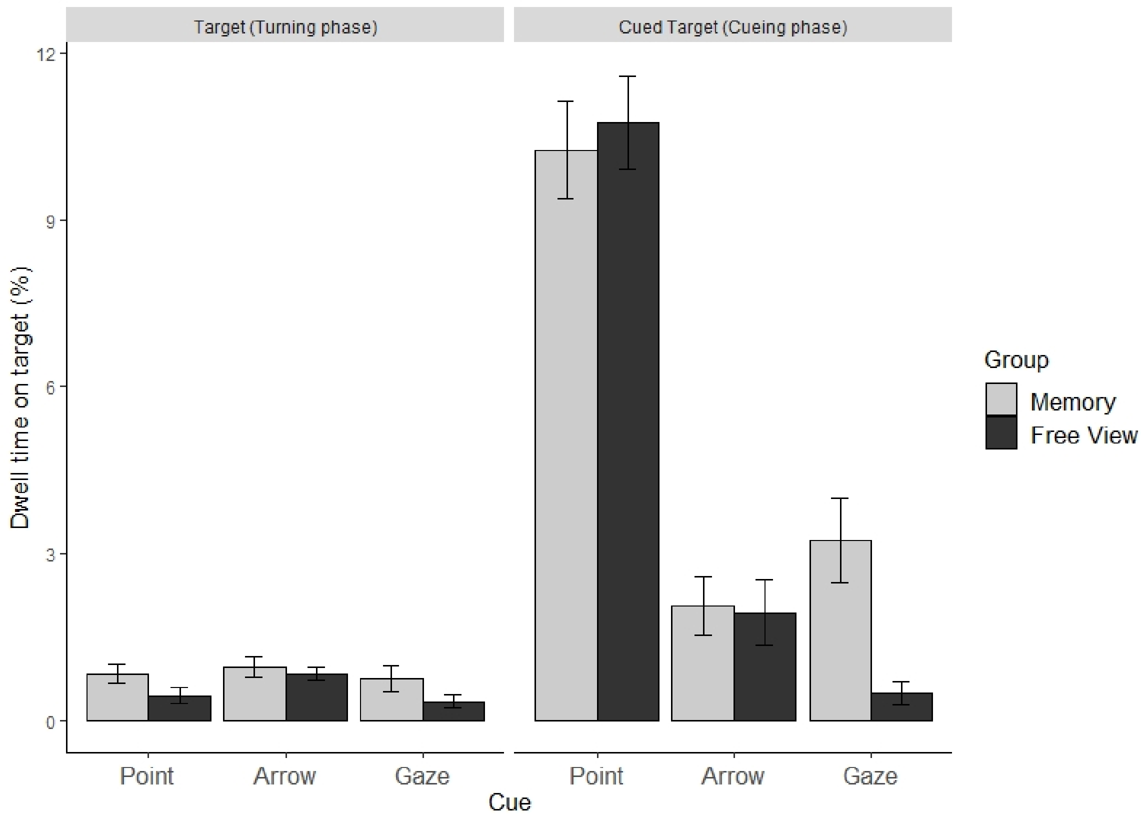

Fig. 3.5 Dwell Times on the target for the three cueing conditions (horizontal axis) for the two motion stages (Turning phase = during turning, Cueing phase = during cueing) and tasks (free viewing and memory). Averages were obtained by first pooling across scenes and then across participants. Error bars show the standard error across participants.

To test whether these observed differences were statistically significant, linear mixed-effects model analyses, using participants and scenes as random factors were conducted. A significant interaction between cue, motion stage and task as found, possibly reflecting the longer dwell times for pointing only within the cueing stage ($\chi^2(2) = 10.12, p < 0.0001$).

Within motion phase analysis

As the patterns of dwell times appear to differ between cueing stages, the effects of task and cue type was first investigated per cueing stage. No interaction between cue and task was found during the turning stage ($\chi^2(2) = 8.76, p = 0.12$). Linear mixed-effects model analysis on the main effect of the cue for the turning stage show no significant differences ($\chi^2(1) = 3.83, p = 0.15$). Similar results were obtained for the main effect of the group in the turning stage ($\chi^2(1) = 3.58, p = 0.06$).

Mixed effect comparison analysis for the cues between the two task groups (turning motion stage) showed significant differences only for the pointing cues in the two groups ($\chi^2(1) = 6.60, p = 0.01$) and no significant differences for the other two cues (arrows:$\chi^2(1) = 0.01, p = 0.94$; gazing: $\chi^2(1) = 1.11, p = 0.29$).

Linear mixed-effects analysis for the interaction between task and cue type was significant during the cueing stage ($\chi^2(5) = 214.11, p < 0.0001$). Linear mixed-effects analysis on the effect of the cues showed significant differences ($\chi^2(2) = 202.97, p < 0.0001$). No significant differences were found for the effect of the group on dwell times on the target.

Mixed effect comparison analysis for the cues between the two task groups showed no significant differences for the arrows ($\chi^2(1) = 0.21, p = 0.65$) and point ($\chi^2(1) = 1.60, p = 0.21$) but a significant difference for gaze cues: $\chi^2(1) = 9.26, p = 0.002$), suggesting that task influenced target dwell times only for the gazing cues.

Within task analysis

Analyses for the memory task during turning stage

Differences between dwell times on the target for the three cues in memory group (turning stage) were also computed. Linear mixed-effects analysis on the differences in dwell times on the target for the three cueing conditions in memory group, showed no significant difference ($\chi^2(2) = 0.78, p = 0.68$). Mixed effect comparison analyses of the cues for the memory group (turning stage) showed no significant differences (arrow vs gaze ($\chi^2(1) = 0.05, p = 0.83$; arrow vs point: $\chi^2(1) = 0.31, p = 0.58$ and gaze vs point ($\chi^2(1) = 0.73, p = 0.39$).

Analyses for the free viewing task during turning stage

Differences between dwell times on the target for the three cues in free viewing group (turning stage) were also computed. Linear mixed-effects model analysis on the differences in dwell times on the target for the three cueing conditions showed no significant differences ($\chi^2(2) = 6.05, p = 0.05$). Mixed effect model comparisons of the cues for the free viewing task showed longer dwell times for the arrows compared to pointing ($\chi^2(1) = 6.28, p = 0.01$) but no differences between the arrows and gaze($\chi^2(1) = 1.64, p = 0.20$) or point and gazing ($\chi^2(1) = 1.41, p = 0.24$).

Analyses for the memory task during cueing stage

Exploring the differences on dwell times between the three cues for the memory task in the cueing stage showed significant differences ($\chi^2(1) = 80.15, p < 0.0001$). Mixed effect comparison analyses exploring the differences between the dwell times on the targets for the three cueing conditions under the memory task showed longer dwell times for the pointing cues compared to gazing ($\chi^2(1) = 46.08, p < 0.0001$) and to arrows ($\chi^2(1) = 57.49, p < 0.0001$) but no differences between arrow and gaze ($\chi^2(1) = 1.47, p = 0.22$).

Analyses for the free viewing task during cueing stage

Similar analyses were conducted for the free viewing task (cueing stage) showing significant differences between dwell times on the targets for the three cueing conditions ($\chi^2(2) = 131.43, p < 0.0001$). Comparing the three cues with linear mixed-effects model analyses result suggest same pattern with the memory task group for the pointing cues (point vs gaze:

$\chi^2(1) = 90.32, p < 0.0001$; point vs arrow: $\chi^2(1) = 62.16, p < 0.0001$), but longer dwell times on the target for the gazing compared to arrows ($\chi^2(1) = 13.84, p = 0.0002$).

3.2.2.5 Trials with fixations on the target

Figure 3.6 shows that in most conditions and motion stages, the target was fixated at least once (percentages of around 70-75%). The percentage of trials with at least one fixation on the target while cue was moving for both groups (Figure 3.6) shows for memory group the largest number of trials with a fixation on the target was only for the gaze cues. For the free viewing task (same motion stage) pointing and arrow cues had the largest number of trials with at least one fixation on the target. During the cueing phase and for the free viewing task, gaze cues shows the largest number of trials with a fixation on the trials on the target followed by arrow cues. Pointing cues show not to be influenced by the task.

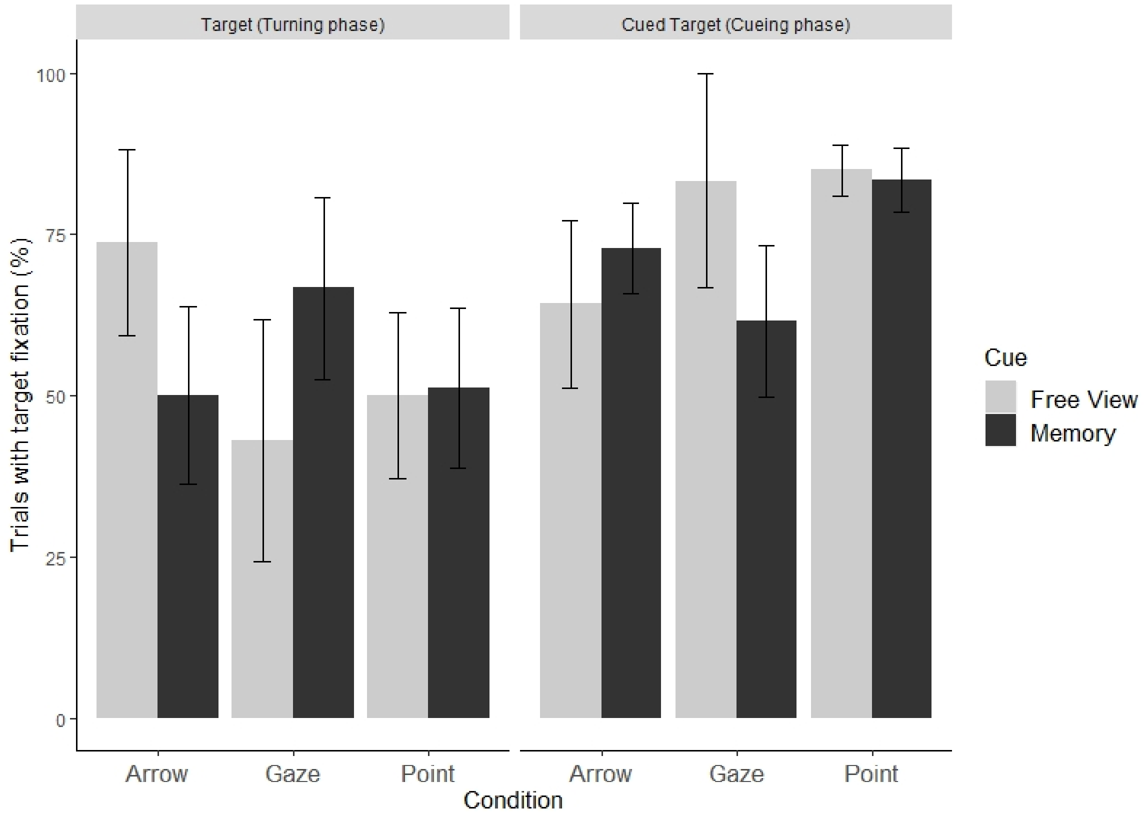

Fig. 3.6 Percentages of trials with at least one fixation on the target for cueing conditions, task groups (free viewing, memory) and motion stages (Target = while cues turning, Cued Target = while cues cueing). Error bars represent the standard error of the mean across participants.

Interaction analysis

A mixed effects logistic regression, using participants and scenes as a random factor, showed no significant three-way interaction between cue type, motion stage and task ($\chi^2(7) = 5.36, p = 0.62$). A mixed effects logistic regression, showed no significant interaction between the two task groups and motion stages ($\chi^2(1) = 1.04, p = 0.61$).

Analysis on the turning motion stage

Mixed effect logistic regression showed no significant interaction between the cueing conditions and the two task groups ($\chi^2(2) = 3.19, p = 0.20$). Linear mixed-effects analysis for the main effect of task groups, on the percentages of trials with at least one fixation on the target, per cueing condition, in turning motion, showed no significant differences for the two social cues (point: $\chi^2(1) = 0.67, p = 0.41$; gaze cues: $\chi^2(1) = 1.81, p = 0.18$) and the arrows ($\chi^2(1) = 0.92, p = 0.34$).

Detailed comparison analysis (linear mixed-effects models) in each task group for the same motion stage showed no difference between the two social cues for the free viewing group (gaze vs point: $\chi^2(1) = 0.76, p = 0.87$; gaze and arrow: $\chi^2(1) = 0.77, p = 0.80$; point vs arrow ($\chi^2(1) = 0.16, p = 0.88$) and no difference between the cues for the memory group (arrow vs point: $\chi^2(1) = 0.45, p = 0.50$; gaze vs arrow: $\chi^2(1) = 1.03, p = 0.31$; gaze vs point: $\chi^2(1) = 0.12, p = 0.73$). There are substantial numbers of trials without a fixation on the target (around 25% for the pointing and arrow cues, and around 40% for the gazing).

Analysis on the cueing motion stage

While on cueing motion stage, linear mixed-effects model analysis for the interaction between the two task groups and the cueing conditions showed no significant interaction ($\chi^2(2) = 4.92, p = 0.09$). Linear mixed-effects analysis on the main effect of the task groups per cueing condition showed no significant differences for any of the cue conditions (gaze: $\chi^2(1) = 0.01, p = 0.93$; arrow: $\chi^2(1) = 0.18, p = 0.67$; and point: $\chi^2(1) = 1.71, p = 0.19$). Detailed comparison analysis using mixed effect logistic regression model for each task group showed no significant differences for any of the cues in the memory task (gaze vs arrow: $\chi^2(1) = 0.09, p = 0.76$; gaze vs point: $\chi^2(1) = 2.28, p = 0.13$; point vs arrow: $\chi^2(1) = 1.02, p = 0.32$) but significant differences only for the arrow compared with other two social cues in the free viewing task (gaze vs arrow: $\chi^2(1) = 5.81, p = 0.02$; point vs arrow: $\chi^2(1) = 4.46, p = 0.04$; gaze vs point: $\chi^2(1) = 011, p = 0.74$). Again there are substantial numbers of trials without a fixation on the target (approximately 20% to 40% for arrows; 40 to 50% for gazing and 25% for pointing).

3.2.2.6 Direction of saccades

The number of trials with fixations on the cue and the target allows for an analysis of the direction of the saccades from the cue (i.e., where do the eyes go after fixating the cue), which provides the most direct measure of how strongly the cue shifts an observer's attention to the target (Hermens & Walker, 2015). Figure 3.7 shows that most saccades, for both task groups, leave the cue, but do not go to the target, suggesting relatively weak cueing. Note here that only the saccades the cueing stage were calculated as it was difficult to determine which section of the turning motion needed to be analysed.

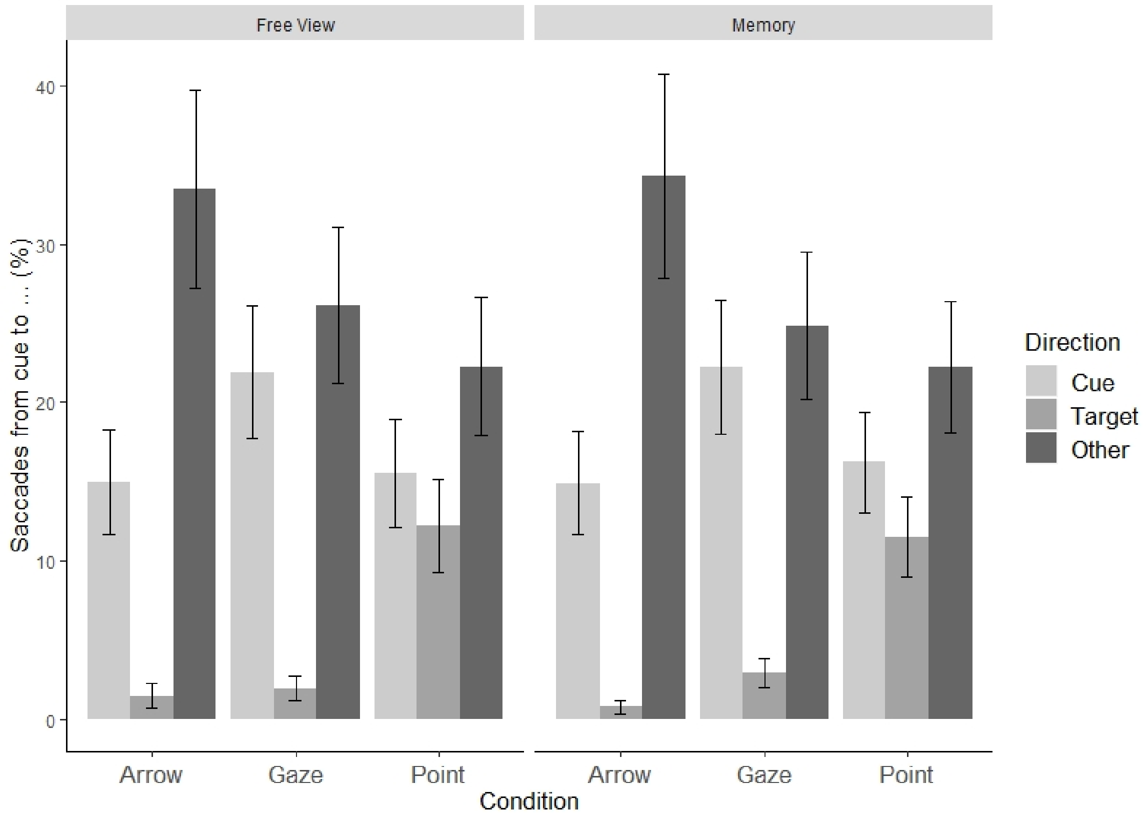

Fig. 3.7 Saccades from the cues to the three directions (target, cue and environment) for the two task groups (free viewing, memory). Error bars represent the standard error of the mean across participants.

Whether just the first fixation on the cue is used, or all fixations on the cue, does not seem to make much of a difference. In line with Hermens and Walker (2015) all saccades are therefore used here. Figure 3.7 suggests that most saccades were directed to a section of the scene that was not the cue or the target (dark bars). For arrows, there were few saccades that go to the target. More saccades were directed to the target for gaze and pointing cues, but these numbers are fairly variable across participants (large between subjects' error bars).

A mixed effects logistic regression showed that the direction of saccades depends on the cue (interaction between cue condition and groups ($\chi^2(12) = 24.83, p = 0.02$) suggesting more saccades leaving the cue and going to the environment. First block of trials analysis showed similar significant results ($\chi^2(12) = 21.6, p = 0.04$). These differences reflect differences only in saccades going back to the target for both groups ($\chi^2(5) = 50.25, p < 0.0001$), but not saccades going back to the cue ($\chi^2(5) = 4.61 p = 0.47$) and saccades going elsewhere ($\chi^2(5) = 5.30, p = 0.38$). First block of trials analyses showed similar results for the saccades leading to the target ($\chi^2(5) = 35.10, p < 0.0001$) as well as for the saccades going back to the cue ($\chi^2(5) = 1.86, p = 0.87$) and elsewhere ($\chi^2(5) = 5.81, p = 0.32$).

If only the saccades leaving the cue and directed to the target considered in relation to the cues and two groups, none of the cues showed an influence from the task (gaze cues: $\chi^2(1) = 0.62, p = 0.43$; arrow cues: $\chi^2(1) = 0.72, p = 0.40$; point cues: $\chi^2(1) = 0.03, p = 0.86$). Similar results were found when only the first block of trials was analysed (gaze cues: $\chi^2(1) = 0.04, p = 0.84$; pointing cues: $\chi^2(1) = 2.05, p = 0.15$ and arrow cues: $\chi^2(1) = 1.039, p = 0.31$).

Exploring in detail the differences between cues for saccades to the target in memory group a linear mixed-effects analysis showed differences between pointing and the other two cues (pointing vs gaze: $\chi^2(1) = 10.73, p = 0.001$; pointing vs arrow: $\chi^2(1) = 16.22, p < 0.0001$) and between gazing and arrow ($\chi^2(1) = 4.60, p = 0.03$). Analyses on the first block of trials showed similar significant results (gaze vs pointing: $\chi^2(1) = 6.94, p = 0.01$; gazing vs arrow: $\chi^2(1) = 2.15, p = 0.01$ and arrow vs pointing: $\chi^2(1) = 9.02, p = 0.003$) Analysis of the saccades going to the target for the free viewing group showed similar pattern with pointing cues leading to the target more than gazing and arrow (pointing vs gazing: $\chi^2(1) = 12.80, p = 0.0004$; pointing vs arrow: $\chi^2(1) = 11.81, p = 0.0006$) and arrow vs gazing ($\chi^2(1) = 0.18, p = 0.67$). First block of trials analysis showed similar results (pointing vs gazing: $\chi^2(1) = 9.85, p = 0.002$; arrow vs gazing: $\chi^2(1) = 0.46, p = 0.50$ and arrow vs pointing: $\chi^2(1) = 9.45, p = 0.002$).

3.2.3 Discussion

The present experiment compared the influence of dynamic social and symbolic cues on the observer's eye movements during free viewing and a memory task. Past research has suggested that when presented with images containing people, observers prioritize actors' faces (Birmingham, Bischof, & Kingstone, 2008a; Fletcher-Watson et al., 2008; Yarbus, 1967). These findings seem to be consistent even in studies where people are part of a natural scene (Birmingham et al., 2009; Zwickel & Võ, 2010; see also the findings in Chapter 2). While people strongly draw the observer's attention, cueing effects from these people are less clear (e.g., Hermens & Walker, 2015; Zwickel & Võ, 2010).

While static social and symbolic cues have been extensively explored, fewer studies have used more realistic dynamic cues (showing, for example, a pointing movement, rather than the result) (see Bayliss et al., 2005; Farroni et al., 2000; Hermens & Walker, 2012; Kuhn & Tipples, 2010; Rutherford & Krysko, 2008; Swettenham et al., 2003). Likely reasons for more common use of static cues is (1) that they are easier to construct (photos rather than videos), (2) that they are easier to present (software to deal with videos as stimuli often does not work as smoothly as for presenting images), and (3) that there may be a confound between the effects of the motion itself and the social aspect of the cue (particularly because imaging studies have suggested that similar brain regions are involved). Despite these possible issues, some studies have used dynamic cues and also found that observers are strongly drawn towards faces of people in the scene (Kuhn et al., 2009). However, it is yet unclear whether the same actors in the scene can also direct the attention of the observer to an object in the scene, and whether this cueing effect depends on the type of cue (social or symbolic). The present experiment was designed to address exactly these questions.

The first aim of the experiment was to determine whether social cues, and particularly gaze cues, were looked at more than symbolic cues (arrows). In order to analyse the gaze patterns, dwell times on cues and targets were split for the different phases of the video clip: (1) the turning motion (or the arrow appearing), (2) the static section after the turning movement (more in line with static cues). The pattern of dwell times on the cues strongly depended on the phase of the video clip. While turning, dwell times on the arrows were longer than on the two social cues. During the more static phase, differences between the cues were much smaller. A possible reason why arrows drew such strong attention during the turning phase is that the arrow appeared during this phase and the appearance of objects has been shown to strongly influence the human oculomotor system (e.g., the remote distractor effect, Walker, Deubel, Schneider, & Findlay, 1997), whereas changing or disappearing objects have much less of an effect (Reuter-Lorenz, Oonk, Barnes, & Hughes, 1995). Dwell times on the arrows while in the static phase, are in agreement with the findings from experiments in Chapter 2, where arrows were presented from the initiation of the trial.

From the two social cues, the pointing cue had longer dwell times than the gaze cue during the turning phase, but only for the free viewing group. The stronger effect of the pointing cue can be understood from the more pronounced motion, and from a more active form of social attention that this cue provides (e.g., Baron-Cohen, 1995). The task effect is interesting and not entirely unexpected from the literature on the effect of task on people's eye movements (e.g., Fletcher-Watson et al., 2008; Itier et al., 2007; Yarbus, 1967), but the exact nature of the task effect is a bit more difficult to understand. Previously, the memory task seemed to lead to less focused viewing of the scenes (Hermens & Walker, 2015), and this may be what is happening here. Still, it is surprising that such effects only seem to occur for the gaze and not for the pointing and arrow cues, and this effect therefore needs to be followed up in future studies.

One confounding element in the dwell times on the different cues while turning is that the movements take different amounts of time (see also Hermens and Walker (2012) for a similar argument). Arrows appeared and therefore seemed to have drawn strong attention, the gaze turn was short and quick and may therefore have led to drawn stronger attention than the relatively longer pointing gesture, but only when participants were not occupied with a task at hand, but this is largely speculation and needs to be addressed in future work.

The cueing stage of the movement (when no movement occurred) was much more like a static cue, and it is therefore important to look back for the interpretation of these results in Chapter 2 in this thesis (particularly Experiment 1 where only one cue was used). In Chapter1, findings showed that observers tend to look at the actors in the scene, and in particularly when they are gazing. In addition, whether the actor was looking at an object or pointing at that same object always showed a stronger effect to capture the observer's attention compared to an arrow.

Similarly, for the dynamic cues, dwell times on the cues during the static phase were longest for the gaze cues, compared to the arrows and pointing cues. These results are consistent with the findings from Chapter 2 and the literature (e.g., Birmingham et al., 2008a, 2009; Fletcher-Watson et al., 2008; Fletcher-Watson, Leekam, Benson, Frank, & Findlay, 2009; Zwickel & Võ, 2010). One potential reason why the pointing cue and the arrows were looked at less was that they were already looked at more during the turning phase and therefore no longer as interesting to the observer. In agreement with studies like Hermens and Walker (2015), dwell times on the cues were not influenced by the task at hand. This is not surprising that at this motion stage, actors and arrows had assumed the same position as they were presented in Hermens and Walker (2015), as well as in Chapter 2 experiments (e.g., actors' full body and face turned to towards the targeted object position).

One contrasting finding with the static cues was that the arm drew attention during the turning phase. For static cues, past work (e.g., Birmingham et al. (2009) as well as findings from experiments in Chapter 2 showed that people's faces attract most of the attention, even while the actor is pointing. When observers see the actual pointing movement, the results presented here therefore suggest that this pattern changes and the arm starts to receive attention. Part of the reason is probably that the moving arm shows quite a bit of motion, which is known to attract attention (e.g., Ludwig, Ranson, & Gilchrist, 2008), but social aspects of the pointing movement are also expected to play a role (references on pointing gestures here).

One of the aims of this study was to compare the cueing effects of the different dynamic cues. There is no clear single eye movement measure that expresses this cueing effect, and therefore various measures were examined: (1) dwell times on the cued object (which are likely to be influenced by whether the cue itself occupies the viewer), (2) trials in which target was looked at or ignored, (3) the number of saccades directed to the target from the cue (argued by Hermens and Walker (2015), to be the most direct measure of cueing in natural scenes).

Intuitively, it may be expected that dynamic cues direct an observer's attention more strongly towards the target, because they are clearer acts to indicate social attention than static indications of objects of interest. Comparing the results of the present experiment with the static cues in Chapter 2, this does not seem to be what happens. Similar to Chapter 2 studies, cueing effect to the target was fairly low, with more saccades leading to other items in the environment. These findings seem to be consistent when embedding cues in natural scene and it seems to be independent from the way cues are presented (static or dynamic). In addition, pointing cues led more successfully to target for groups (memory and free viewing).

These results are at odds with previous findings. Previous studies using static cues have found a strong cueing effects (Driver et al., 1999; Friesen et al., 2005; Langton et al., 1996, 2000; Ricciardelli, Betta, Pruner, & Turatto, 2009; Ricciardelli et al., 2002). The weak cueing by gaze cues in particular was also at odds with studies that showed dynamic gaze cues in isolation and at fixation (e.g., Bayliss et al., 2005; Farroni et al., 2000; Kuhn & Tipples, 2010; Ricciardelli et al., 2002; Rutherford & Krysko, 2008; Swettenham et al., 2003). It therefore seems that either embedding the cues in a natural scene, or by not presenting them directly at fixation (Hermens et al., 2015), they lose their potency in directing an observer's gaze. A possible reason why cues presented in isolation and at fixation have a stronger effect is that they may lead to a learning effect (e.g., Driver et al., 1999; Friesen et al., 2004), where participants learn that the cues are important for the study. The lack of a background may also place extra focus on the cues, which is not observed for cues embedded in a natural scene. A further complication of cues in a natural scene is that the direction of a cue is much more difficult to judge due to distortions arising from the projection of a 3D scene on a 2D retina (Doumen et al., 2010; Hermens & Gielen, 2003) and the variety of items surrounding the cued item. An additional reason for the cueing effect in the current study, is that the dynamic cues drew and preoccupied the observers' attention for longer, leaving less time to look at the target objects.

For the arrows an alternative explanation can be offered for their failure to show a strong cueing effect. Due to the way arrows were presented their cueing effect might have pre-existed before the time this study examined. As it is difficult to determine at which point during the turning stage, social cues shifted attention to the target, this question cannot be answered at the moment and further investigation is needed. Task did not influence the cueing effect for any of the cues. This comes as a surprise as previous studies have showed a task influence on cues' attentional shifting (e.g., Hermens & Walker, 2015). Potentially these results are due to the movement proceeded the cueing stage. The absence of movement might have led the observers in both groups to show an equal interest to follow where the cues (specifically social cues) looking or pointing at. However, this is merely a speculation and in the future a better balance between motion and task is need to distinguish the source of the observed effects.

3.3 Conclusion

The present study examined the influence of dynamic social and symbolic cues on observers' attention. The results show that gaze cues attract participants attention, but only when actors have assumed their cueing position. In contrast, pointing cues led to the strongest cueing, but cueing was generally weaker than expected on the basis of past work. Motion seems to influence observers' time looking at the cues and targeted items. Overall task did not affected observers' attention towards the cues and this might be derived from cues' motion.

3.4 From the screen to the real-world

The two chapters so far have examined cues that were presented on a computer screen. While presenting cues in such a setting is convenient, screen presentation is unlike how people normally encounter such cues. In fact, previous work that directly compared people shown on a screen or in the real-world demonstrated that observers shy away from looking at a real person, while staring at a person on a screen (Laidlaw et al., 2011), possibly because looking at real people may lead to engagement with the person, or impressions of being rude. While this study looked at gaze behaviour towards other people, it has not examined how observers respond to direction cues from these other people, such as pointing gestures or head turn. Chapter 4 therefore moves from the lab into the real-world and examines how gaze cues affect attention in this setting.

Chapter 4

Exploring social and symbolic cues in the real-world

4.1 Introduction

A common activity in day to day life is socializing with other people. During such interactions, it is not uncommon to take notice of other people's facial expressions and gestures. When moving around, people also encounter symbolic cues that indicate a direction (often arrow signs). An ongoing debate has been whether cues of direction provided by others (e.g., head turns, pointing gestures) provide stronger cues towards objects of interest than symbolic cues, such as arrows. It may be argued that there is an evolutionary advantage for social cues (as arrows have not been introduced until fairly recently), and that within an individual life-span (in particular during childhood) there is a stronger exposure to social direction cues than symbolic cues.

The majority of the literature is in line with such arguments, showing that other people strongly attract attention of observers, and that images of gazing faces or pointing hands more strongly direct an observers' attention than arrows (e.g.,Althoff & Cohen, 1999; Firestone, Turk-Browne, & Ryan, 2007; Henderson, Williams, & Falk, 2005; Itier et al., 2007; Schyns, Bonnar, & Gosselin, 2002; Vinette, Gosselin, & Schyns, 2004; Yarbus, 1967). A drawback of this literature is that it strongly depends on a rather impoverished version of the real-world. In a typical social cueing experiment, an image of a face with the eyes turned left or right is presented at fixation in an otherwise empty display (followed by a peripheral target to probe into attention effects). An important question therefore is whether this impoverished version of a social cue has similar effects on observers as actual social cues encountered in the real-world. Recent studies have suggested they may not. For example, Hermens et al. (2015) found that gaze cues (eyes turned left or right) were not effective when they were not (initially) presented at fixation. Moreover, in a real-world setting, observers avoided looking at people, while they still looked at a face that could be seen on a computer display (Laidlaw et al., 2011).

In addition more recent studies (e.g., Gregory & Antolin, 2019) have revealed that avoiding looking at other people's faces is not only related to the way gaze cues are presented (e.g., real person vs person in the screen). Instead, observers tend to avoid people in the screen when there is a live stream interaction compared to when they watch a pre-recorded person. This suggests that the mere chance for an interaction with an unknown person initiates an avoidance response towards the face of the confederate in the screen. However, a study by Freeth, Foulsham, and Kingstone (2013) have suggested that people spend more time looking at another person's eyes, face and body in real life interaction and their gaze behaviour is influenced by the confederate eye direction. These results were mainly observed only when participants were asked a question. This suggests that looking at another person occurs only under certain scenarios and when the other person is essential to our understanding of the situation (e.g., using moving speech and body language to decode speech). Potentially when interaction is not direct attention on the face is not necessary and avoidance response comes into play.

Studies have also examined whether the background on which the cues are presented (natural scene or empty display) and whether the entire body is visible rather than just the head or hand, has an effect (e.g.,Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010; and the studies in Chapter 2 and 3). In such settings, actors in the scene were looked at more than arrows (Birmingham et al., 2009; Hermens & Walker, 2015) or standing loud-speakers (Zwickel & Võ, 2010). Cueing (as measured by the direction of saccades leaving the cue), however, did not seem to differ as much between the cues, with a slight advantage for actors that were pointing rather than gazing (or arrows pointing).

There are important reasons why studies have typically relied on images or video clips rather than real-world gaze cueing: (1) various aspects (e.g., size, contrast) of the stimuli can be much better controlled for images (and to some extent videos), (2) stimuli can easily be presented and repeated, (3) no real-world actors needed for each participant tested. There are also important reasons to move such studies to the real-world. First, as indicated above, it has been shown that looking at people in the real-world people is often avoided (Gallup, Chong, & Couzin, 2012; Laidlaw et al., 2011), possibly because looking at a person is likely to initiate an (unwanted) interaction. Research in pedestrians has also suggested that following gaze of others is avoided when that person approaches the viewer (Gallup, Chong, & Couzin, 2012; Gallup, Hale, et al., 2012), but not when the viewer sees the person from behind. Other studies comparing laboratory and real-world settings, also found differences in eye movement patterns (Land & Tatler, 2009). Second, studies with images on a computer screen do not embed the viewer into the scene and take away most of the 3D aspect of the real-world.

Studies of visual attention often ask participants to ‘simply look’ (free viewing). In day to day vision, however, such viewing is not as common, and often viewers look for a particular reason (e.g., to find an object, to go to a particular location). Past studies have suggested that eye movements can vary strongly dependent on the task that observers are performing (e.g.,Dukewich et al., 2008; Nummenmaa & Hietanen, 2009; Yarbus, 1967), although not all tasks lead to differences in eye movements (e.g.,Hermens et al., 2018; Hermens & Walker, 2015; Chapter 3 of this thesis). Social cues may be particularly important when looking for a search target, and cues may work especially well in a real-world setting, where there would be less reason to believe the actor was not simply being helpful.

Natural environments often use signs of directional cues, such as arrows, but also pointing hands. It is unclear how effective such cues are in a real-world task. It may be expected that if actual people provide the cues, they may be more effective than signs with such cues. On the other hand, approaching pedestrians that looked at a target object were not often followed (Gallup, Chong, & Couzin, 2012; Gallup, Hale, et al., 2012) and it is therefore important to compare these two types of cues directly in a controlled experiment. With this information in mind, this chapter aims to explore how direction cues on signs and direction cues provided by actual people influence gaze behaviour in an observer. Specifically, the study aimed to test (1) whether schematic faces (presented in a sign) have weaker or stronger cueing effects than actual people in a real-world scene, and (2) whether prior knowledge of the search object aids localizing the object, and (3) whether cues aid localizing the search target and (4) whether gaze cues can capture the observer's attention. Two real-world visual search studies were conducted. In Experiment 5, participants were asked to locate four common items each placed in a different room. Signs with (1) a pair of eyes, (2) a pointing hand, and (3) an arrow were used to indicate the location of these objects, whereas the fourth object was paired with no cue. The pairing of the cue with the room and therefore search object was systematically varied across observers, and observers only saw each room once (to avoid memory effects). Using a mobile tracker, participants' gaze behaviour was recorded while they searched for the items.

In a second experiment, actors rather than signs (apart from the arrow) were used. This study also tested whether knowing about the search targets had an influence on how often the cue was viewed and examined whether additional hints during search influenced gaze patterns. Participants were asked to search for two rather than one item inside the same room. Items' number was minimized due to space limitation and to avoid familiarity with the room from multiple visits. Social cues were provided by actors (who happened to work in the rooms used). As in the first experiment, gaze behaviour was recorded with a mobile eye tracker.

4.2 Experiment 5

4.2.1 Methods

4.2.1.1 Participants

Thirty-five participants (20 males and 15 females) took part in the study. Participants were a mixture of university staff, members of the public as well as undergraduate and postgraduate students and their age was between 19 to 68 years (Mean=32.83, SD=13.19). Their vision was either normal or corrected-to-normal with contact lenses (the eye tracker could not cope with glasses). Participants were recruited using an online recruitment system (SONA system), posters and by word of mouth. Those who were recruited using the poster or word to mouth each received £5 for their time, whereas the others were reimbursed with course credits. All participants were naïve to the task and provided written consent for their participation that was approved by the School of Psychology, University Lincoln, ethics committee.

4.2.1.2 Apparatus

Eye movements were recorded with a Tobii 2 eye tracker. The Tobii 2 glasses consist of two units: a headset and a recording unit. The head gear consists of a head centered video camera which records video image from participants' perspective at a frame rate of 25Hz. The field of view of this camera was 82° horizontally and 52° vertically and was fixed to the head due to the fixed position of the camera. Participants' gaze position was recorded at 50Hz by four cameras placed at the inner lower parts of the head gear's frame (two cameras under each eye). Because of the difference between the sampling rate of the camera (25Hz) and eye movement recordings (50Hz), the eye movement data were down-sampled to 25Hz. Gaze direction was estimated from IR images of the participants' eyes, from which the pupil and corneal reflection position were extracted and combined to create a 3D model of both eyes. Calibration of the system involved participants fixating a single calibration target positioned at approximately 1 to 1.5m (a little over an arm's length). The recording unit was attached to participants' clothes or carried in a backpack and collected and stored videos recordings from the scene camera as well as the eye movements on a SD flash card that could be read out using a standard card reader.

4.2.1.3 Stimuli

Each participant entered four different rooms varying in size and the amount of visual clutter, located on the University of Lincoln campus. Rooms consisted of two cubicle offices (small rooms), one office (medium size) and a kitchen (large size) - see Figure 4.1 for images of the different rooms. Before entering each room, participants were provided with a black and white A4 size print of the target item (a cup for the kitchen, a pen for cubicle 1, correction fluid for cubicle 2, and tape for the office). Items were common items for the rooms and placed at locations where they may be expected (e.g., pencil on a desk, mug on the kitchen counter) and were surrounded by various other items in their normal positions.

4.2.1.4 Design

Participants each entered the rooms in the same order - starting from the first cubicle and finishing at the kitchen. A A4 sheets placed in three of the rooms showed one of three cues (cartoon eyes gaze, cartoon hand pointing, and arrow sign) indicated the location of the target object (Figure 4.1). The remaining room had no such cue. Gazing eyes were occasionally positioned vertically (see Figure 4.1) and only when there was no other way to position them (i.e., when they were positioned directly above the targeted items). The association between the different cues and rooms was counterbalanced between participants to avoid effects of the room or the search target on the average data.

Fig. 4.1 Screenshots of the four different cues and rooms used in Experiment 5. Upper left image shows the office with the pointing cue. Upper right image shows the kitchen with the arrow cue. Left (down) image shows one of the cubicles with gazing cue. Right (down) image shows a different cubicle with no cue).

4.2.1.5 Procedure

Before taking part, participants signed a consent form. They were then fitted with the Tobii 2 glasses and the one-point calibration procedure was performed, as recommended by the supplier of the glasses. The eye tracking was started, and participants were led from one room after another by the experimenter, who supplied participants with a photograph of the target image (A4, black and white) and watched that participants found the correct object. When participants thought they had found the target, they were asked to point at the item and show it to the experimenter. After the completion of the experiment, which took around 1 minute, participants were informed about the experiment, were asked a few questions about what they noticed in the experiment, (e.g., if they noticed the presence of cues), and were given the opportunity to ask questions themselves. All participants reported that they did notice the cues. Time restriction was based on another real-world study Tatler and Tatler (2013); was used to control experiment's duration and to see if participants will depend more on the cues to locate the items.

4.2.1.6 Data Analysis

Horizontal and vertical gaze location for each sample was extracted from the JSON file that coded gaze position along with other information from the eye tracker with the use of a custom-built Perl script. These positions were then combined with the video images from scene camera using a Matlab script. For each room and participant, the timing of the following three events were manually coded (frame by frame): (1) the frame in which the contents of the room was first visible (start of the search), (2) the frame in which the participant looked at the target for the first time (determining visual search times), (3) the frame in which they first touched or pointed at the object (determining manual search times) and (4) the frame in which they looked at the cues until the last frame stopped looking at cues. Additionally, it was coded whether participants looked at the cues or not; whether looking at the cues was depended on the room they were presented; whether they looked at the cues before they located the target and whether search time differ between the targeted items. To ensure coding validity, the videos were coded by a second coder, and comparison of the two data led to a 75.5% agreement and 'substantial' Cohen's k of 0.704 (McHugh, 2012). In case of a disagreement, mutual agreement was sought between the two coders. Data were analysed in several ways. For statistical comparisons, linear mixed-effects models (R package lme4) were used with participants as random variables. For the analysis of percentages, the linear mixed-effects analysis was complemented with Chi-square proportion tests.

4.2.2 Results

4.2.2.1 Looking at the cues

The percentage in which the cues were fixated or not is shown in Figure 4.2, suggesting that cues were mostly ignored. This was confirmed with a χ^2 test of independence ($\chi^2(1) =$

$5.49, p = 0.02$). When splitting by cue type (Table 4.1), participants more often ignored the cue than looked at it particularly the arrow cue.

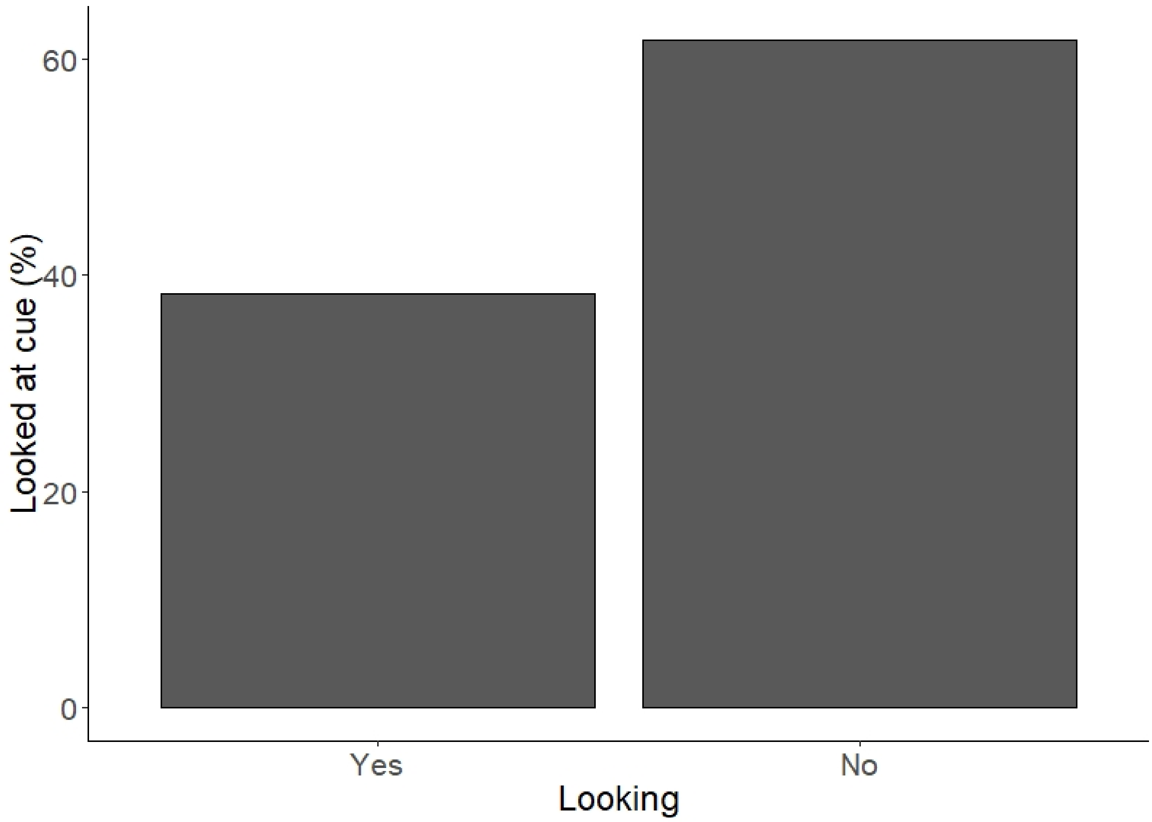

Fig. 4.2 Percentage of trials with at least one frame with a fixation on the cue - across cues.

Three χ^2 tests of independence exploring the differences between the percentage of trials with at least one fixation on the cue (per cue type) showed that arrows were more often ignored than looked at ($\chi^2(1) = 9.26, p = 0.004$), but gaze ($\chi^2(1) = 0.26, p = 0.61$) and pointing cues ($\chi^2(1) = 0.028, p = 0.87$) were equally often ignored to looked at.

Table 4.1 Percentages of trials with at least one fixation on the cue - per cue type.

Cue	Looked at	Percentage
Arrow	Yes	5.88%
	No	27.45%
Point	Yes	14.71%
	No	18.63%
Gaze	Yes	10.78%
	No	22.55%

Figure 4.3 shows the dwell times on the cues for the trials that they looked at the cues. When cues were looked at, they spend more time looking at the pointing, followed by gazing and arrows.

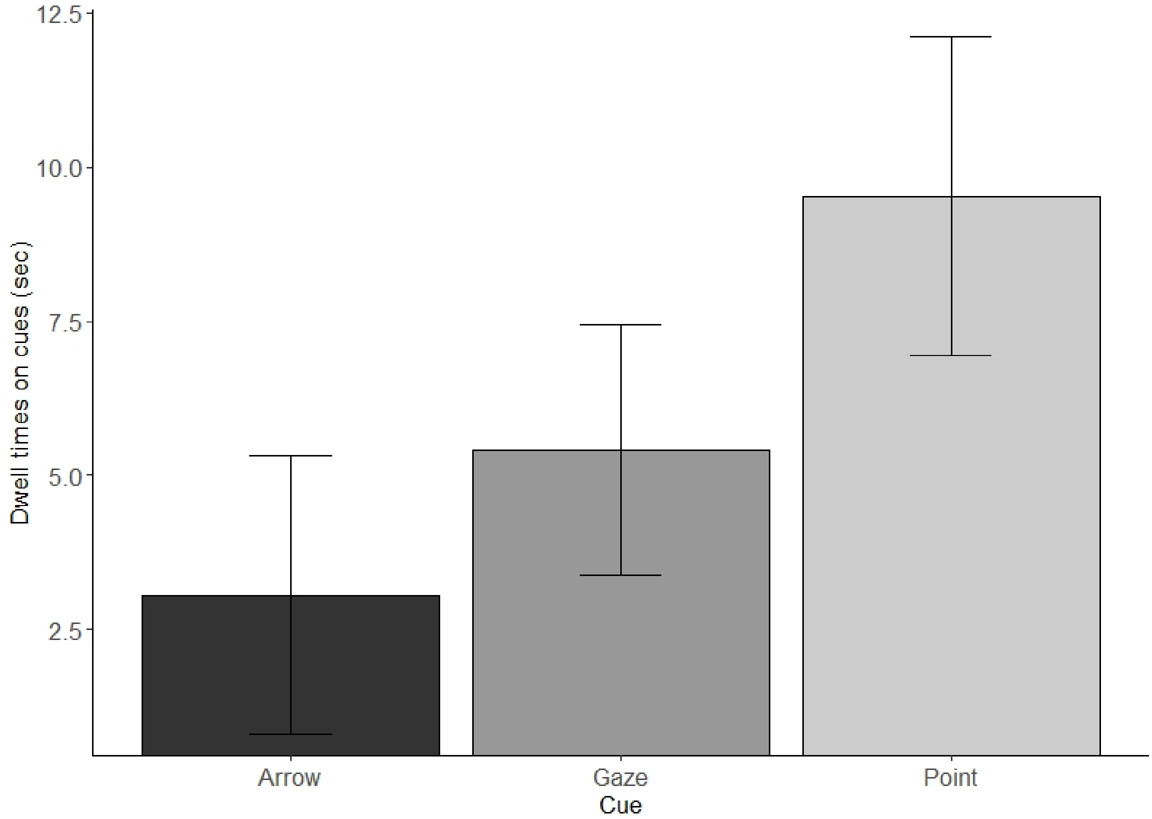

Fig. 4.3 Total looking time on the three cues per second. Averages were obtained by pooling across cues. Error bars represent the standard error of the mean across participants.

To determine whether dwell times on the cues differed significantly, a linear mixed-effects model (with participants and rooms as random factors) analysis was conducted. Linear mixed-effects model showed no significant differences between the cues ($\chi^2(2) = 2.18, p = 0.33$). Linear mixed-effects model analysis comparing the three cues showed no significant difference for the comparison between gaze and the other two cues (gaze vs point: $\chi^2(1) = 2.33, p = 0.13$; gaze vs arrow: $\chi^2(1) = 0.20, p = 0.65$) and a marginal significance between the point and arrow ($\chi^2(1) = 4.02, p = 0.05$). These results suggest that people do not often look at the cues, and when they look at them, they spend a limited time looking at the cue, with no differences between different cues.

Figure 4.4 examines whether noticing the cues depended on the room they were presented. The plot suggests that the cues in the office (large, cluttered room) was more difficult to localize.

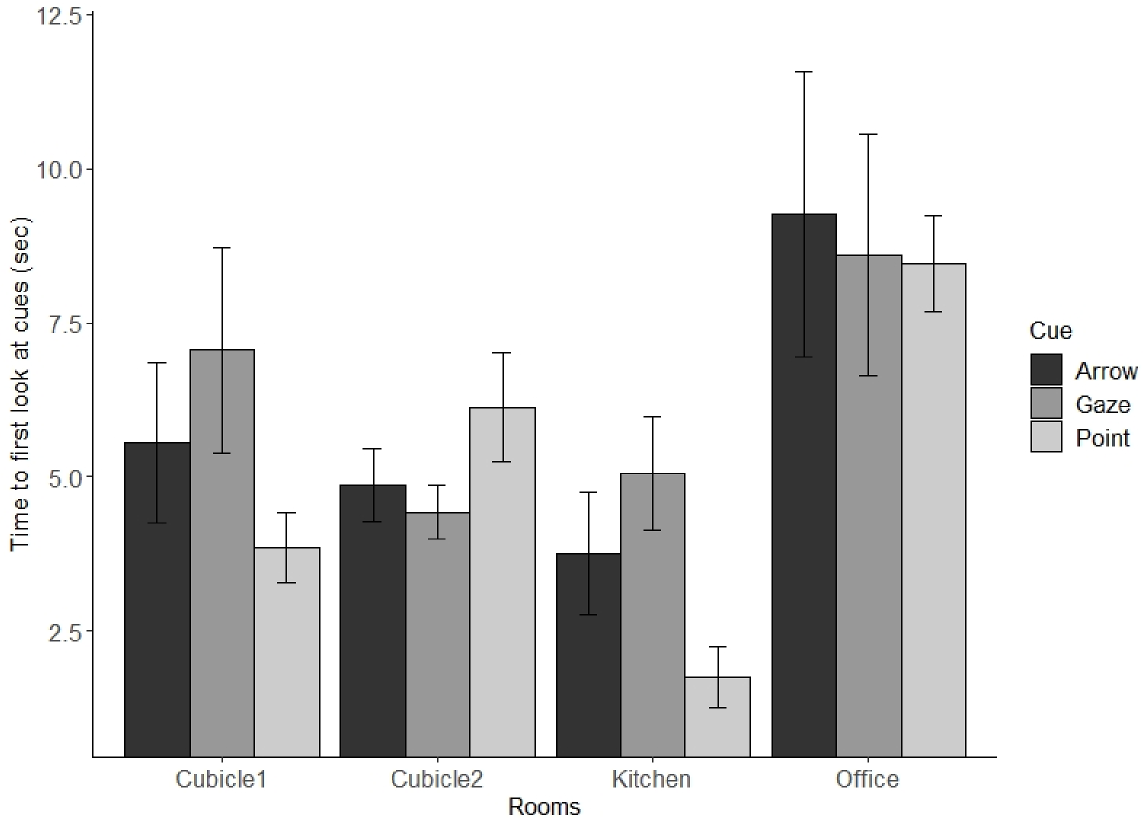

Fig. 4.4 Total looking time on the cues per room. Averages were obtained by pooling across the cues and the four different rooms. Error bars show the standard error of the mean across participants.

Despite these observations, no significant interaction was found between room (and target) and cue type ($\chi^2(6) = 7.01, p = 0.32$). There was a significant main effect of the room ($\chi^2(3) = 24.67, p < 0.0001$), but no main effect of the cue ($\chi^2(2) = 0.97, p = 0.62$). Linear mixed-effects analysis on the main effect of the cue for each room showed no significant difference for the cues in the three rooms (Cubicle1: $\chi^2(2) = 3.21, p = 0.20$; Cubicle2: $\chi^2(2) = 3.80, p = 0.15$ and Office: $\chi^2(2) = 0.15, p = 0.93$) and a marginal significance for

Kitchen ($\chi^2(2) = 6.09, p = 0.05$). Results of paired comparisons comparing the individual cues in each room are shown in Table (4.2).

Table 4.2 Statistics (χ^2-value, p-value) for comparisons of looking time on the cues across the four different rooms (Cubicle 1, Cubicle 2, Office and Kitchen)

Cue	Room	χ^2-value	p-value
Arrow vs Gaze	Cubicle 1	1.82	0.18
Gaze vs Point		2.86	0.10
Arrow vs Point		0.45	0.50
Arrow vs Gaze	Cubicle 2	0.41	0.52
Gaze vs Point		3.12	0.08
Arrow vs Point		1.56	0.21
Arrow vs Gaze	Office	0.10	0.82
Gaze vs Point		0.01	0.94
Arrow vs Point		0.13	0.71
Arrow vs Gaze	Kitchen	0.94	0.33
Gaze vs Point		7.95	0.004
Arrow vs Point		2.60	0.11

4.2.2.2 Finding the target

Figure 4.5 shows how long participants took to find the search target. It suggests that participants were faster to locate the target when it was cued (compared to the no cue condition), but without differences between cues.

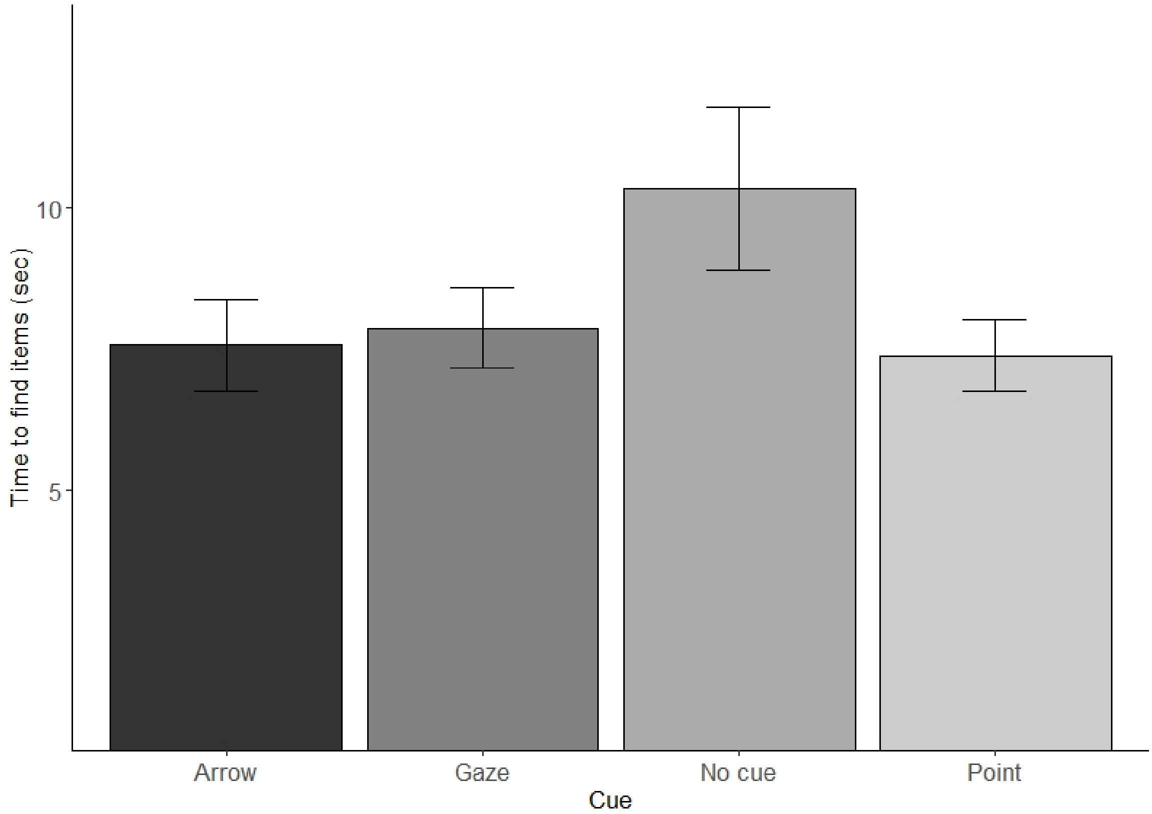

Fig. 4.5 Total time to locate targeted items per cueing condition. Error bars show the standard error of the mean across participants.

Despite the slightly longer search times without a cue, there was no significant effect of cueing condition on search times ($\chi^2(3) = 7.39, p = 0.06$). Pairwise comparisons between search times with a certain cue and the no cue condition showed only significantly faster search times for the pointing cue compared to the no cue condition ($\chi^2(1) = 4.19, p = 0.04$)

Figure 4.6 examines whether cue fixations led to faster localization of the target. For arrows and pointing cues, looking at the cue is associated with slightly faster localization.

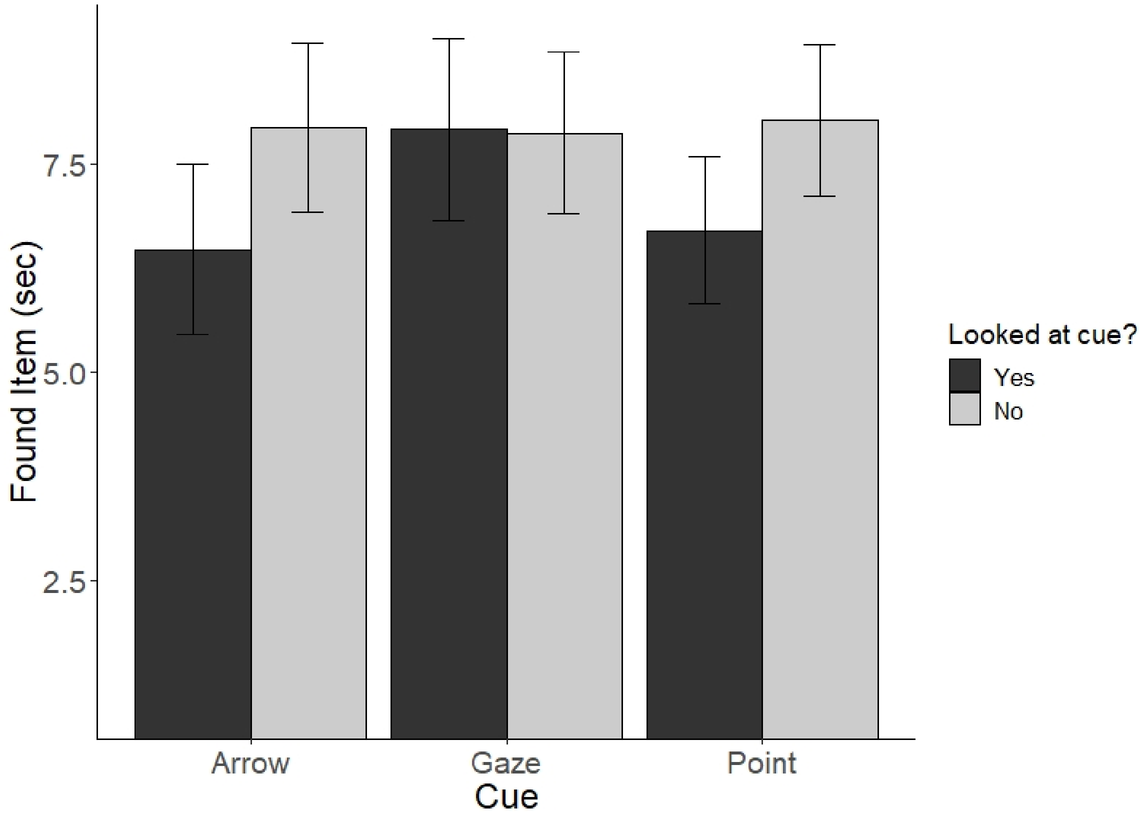

Fig. 4.6 Searching time when looking or not looking at the cue shown separately for the three cueing conditions. Error bars represent the standard error of the mean across participants.

No significant interaction was found between cue type and whether or not participants looked at the cue ($\chi^2(5) = 2.06, p = 0.84$). Linear mixed-effects model analysis exploring the main effect of the cues for the two looking conditions showed no significant differences (Looked at cues: $\chi^2(2) = 1.21, p = 0.55$; Did not look at the cue: $\chi^2(2) = 0.1, p = 0.99$). Subsequent comparison analyses for the cues in the Looked at cues condition, showed no significant differences (Arrow vs Gaze: $\chi^2(1) = 1.68, p = 0.20$; Arrow vs Point: $\chi^2(1) = 0.01, p = 0.97$; Gaze vs Point: $\chi^2(1) = 0.80, p = 0.37$). No significant results were found also for the cues in the conditions where cues were not looked (Arrow vs Gaze: $\chi^2(1) = 0.02, p = 0.97$; Arrow vs Point: $\chi^2(1) = 0.19, p = 0.67$; Gaze vs Point: $\chi^2(1) = 0.03, p = 0.86$). These results suggest that looking or not the cues did not influence participants performance to locate the target.

Finally, analyses were carried out to explore if there were any differences between locating the four items. Figure 4.7 shows that participants required more time to locate the tape, following by the pencil, cup and finally the correction fluid.

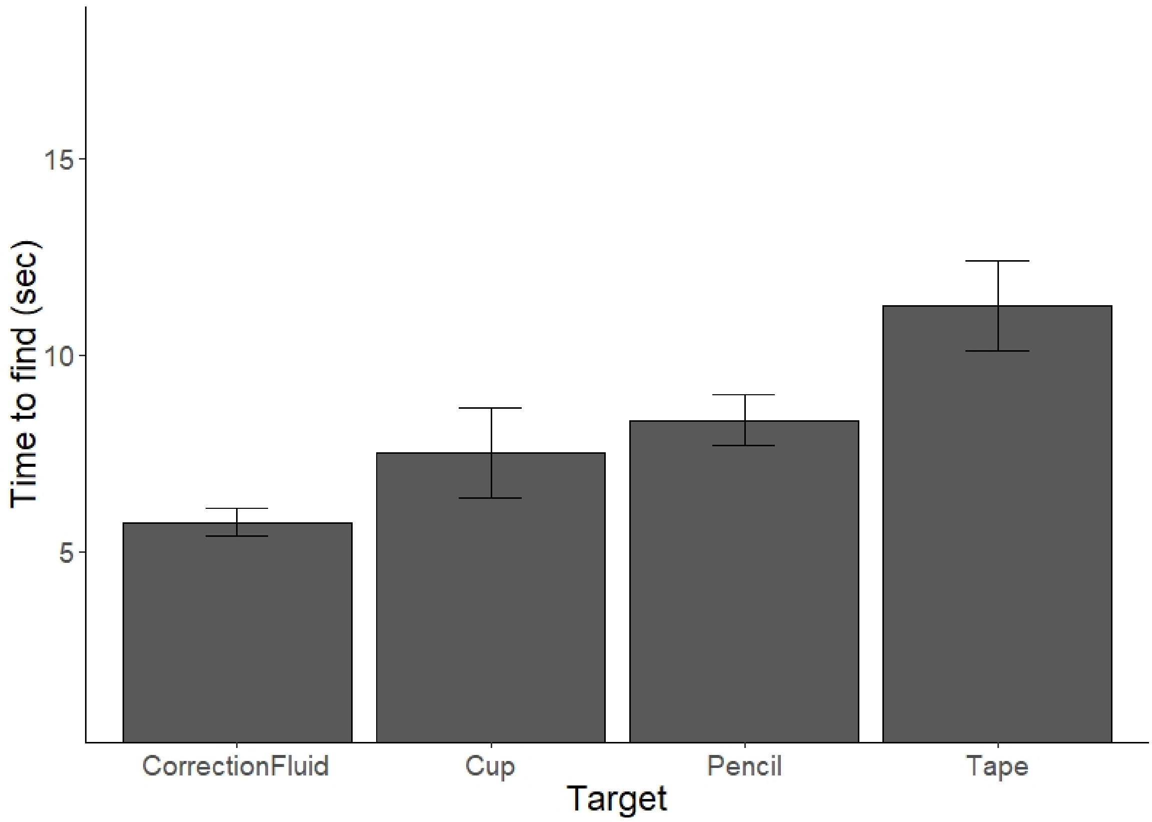

Fig. 4.7 Total Search time per item. Averages were obtained by pooling across the four different items. Error bars show the standard error of the mean across participants.

Linear mixed-effects models, with participants as a random factor, showed significant differences ($\chi^2(3) = 19.31, p = 0.0002$). Results of paired comparisons of looking time on the four different item are shown in Table (4.3).

Table 4.3 Statistics (χ^2-value, p-value) for comparisons of looking time on the four different items (Correction Fluid, Cup, Pencil and Tape)

Target1	Target2	χ^2-value	p-value
Correction Fluid	Pencil	20.79	>0.0001
Correction Fluid	Cup	2.55	0.11
Correction Fluid	Tape	18.90	>0.0001
Cup	Tape	5.27	0.02
Cup	Pencil	0.41	0.52
Tape	Pencil	4.88	0.03

4.3 Experiment 6

Experiment 5 explored the effects of static cues in a real-world setting. Participants looked at the cue in about half of the trials but looking or not at the cue did not lead to faster localization of the target, suggesting that signs with cues are not very effective in attracting or directing the observer's attention. In the second experiment, the effects of actual actors are investigated. Additionally, the experiment will test the effects of task of the observer.

4.3.1 Methods

4.3.1.1 Participants

Initially, 66 participants took part in the study. After filtering for poor quality eye tracking data, 60 participants remained (20 males and 40 females) for the final analysis. These remaining participants were undergraduate and postgraduate students and university staff members (University of Lincoln, UK). Their age was between 18 and 38 years of age (Mean=22.82, SD=4.52) and all had either normal or corrected-to-normal vision. Participants were recruited using an online recruitment system (SONA systems), word to mouth or a recruitment poster. Participants recruited using SONA systems received course credits for their time, while the others received £5 as a reimbursement. Participants were randomly allocated to one of two groups (group 1: received hints that there were cues; group 2: no hints about the cues). This study was approved by the School of Psychology, University of Lincoln (UK), ethics committee. All participants provided consent for their participation.

4.3.1.2 Apparatus

Participants' eye movements were recorded using a Positive Science LLC mobile eye tracker (New York, NY). The right eye (pupil location and corneal reflection) was tracked at a rate of 30Hz with spatial accuracy of 1o. A scene camera at the right side of the frame recorded a video from the point of view of the participants. Eye movement data were stored in real time on an iMac laptop, attached to a backpack, using the Live Capture software, provided by Positive Science. A microphone was also attached to the frame, recording throughout the experiment. These voice recordings were not used for the analysis and therefore -to protect the privacy of the participants in the stored data - excluded from the final videos' rendering. In order to calibrate the eye tracker, participants are asked to fixate a single target and move their head in different directions. Calibration was conducted twice at two different distances (i.e., 167 cm and 321 cm) to minimize the parallax error caused by single tracking eye camera which could have resulted to poor results. The eye position and video recordings during this calibration stage were used to estimate gaze direction was estimated off-line with use of Yarbus software provided by the same company. Calibration was successful for most of the participants, except for in 6 participants, whose data were excluded from the analysis.

4.3.1.3 Stimuli

In contrast to Experiment 5, participants only entered a single room (the technicians' workshop of the psychology department), which helped to control for the spatial layout of the room across cues (See Figure 4.8). The room contained items that varied in size, colour and purpose, and several of the objects were similar to the target object, making search more difficult. Only the three cues were used (i.e., no cue condition was no longer used). The target items (for pictures see Appendix A) used in this study were a tool box and a power supply machine, placed in two different locations inside the room. In contrast to Experiment 5, participants were instructed to search for the target by the experimenter just giving the name of the object (i.e., 'tool box'). This was done in an attempt to make search more difficult and thereby the cues more informative. The two experimental groups differed in how much the experimenter told them about the cues: group 1 was not informed about the cues, whereas group 2 was. It was expected that group 2 would make stronger use of the cues. Both groups were informed that actors' will be inside the room. For Group 1 (not informed group) actors' presence was justified as being there for safety reasons. Group 2 (informed group) was informed that actors are part of the experiment and can help them locate the target. Items were placed at a location not visible to the participant when entering the workshop. Location of the items were not the same, but they always positioned at the same location for every participant. Second item was added before participants entered the room for the second time.

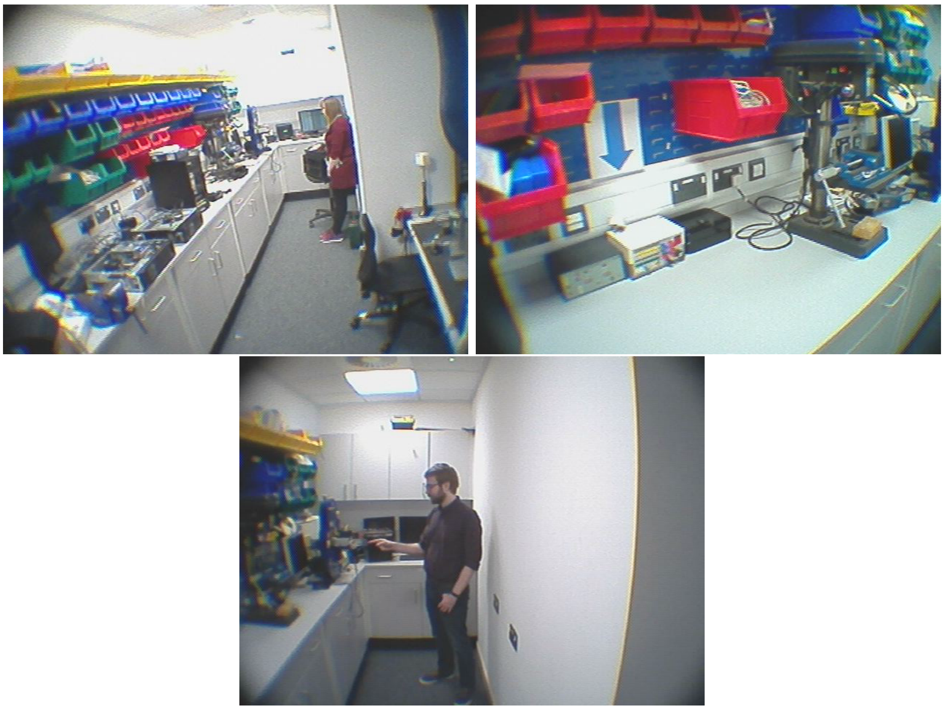

Fig. 4.8 Screenshots from the three cues used in Experiment 6. Left (up) image shows a person gazing at the direction of the second item. Right (up) image shows an arrow pointing at the location of the first item. Left (down) image shows a person pointing at the direction of the second item.

4.3.1.4 Design

This experiment was a between-subjects design, with one factor with three levels (the three different cues). Cues were either social or symbolic. The social cues were provided by an actor (one of the three school of psychology technicians) and involved a pointing gesture or a gaze shift towards the targeted object. Actors' position was approximately opposite to targeted objects' location they were cueing. The symbolic cue was an arrow pointing at the location of the item, already in the room when the participant entered. Cue order was systematically varied between participants and within groups using a Latin square design.

4.3.1.5 Procedure

Before taking part participants signed a consent form. They were then fitted the Positive Science LLC mobile eye tracker glasses and the calibration of the eye tracker was performed. Written and verbal instructions were provided, and the experimenter made sure that the instructions were clear. Participants in the informed group were made aware that there would be helpful hints inside the room to help them locate the items and that the technicians were part of this help. The other group (not informed group) did not receive this information. Both groups were asked not to ask technicians any questions (i.e., items location), not to touch the items and not to open the cupboards (target were not in a cupboard). Each trial started with the experimenter saying out loud the name of the item and leading the participants inside the technicians' office. There one of the technicians escorted the participants inside the room and informing them that they had 1 minute to locate the item. This was done to avoid any surprise from seeing the technicians inside the room that might have caused any unwanted effects on the experiment. This also gave enough time for the technicians to assume their cueing position. Similar to Experiment 5, time limitation was used for several reasons (e.g., control the duration of the experiment). More importantly, time constraint was essential for the current experiment, as participants visited the same room twice. Therefore, time limit aimed to minimize any memory effect from visiting the room twice. Participants were also made aware that they were free to stop looking when they were sure that they had located the target object. When this happened, or when the 1 minute was over, the experimenter came into the room and asked the participants to show the targeted object by touching it. Participants were given one chance only. To provide feedback about their choice, the experimenter led the participants outside the room and showed them a photo of the target object. At the end of the experiment, participants were debriefed and were asked whether they noticed the cues. The majority of the participants reported that they did notice the cues (70%) and that technicians were there to "trick" them and not to help them (63% of the participants in both groups).

4.3.1.6 Data Analysis

Eye movement data were manually coded (frame by frame) and labelled by a single coder. Fixations were categorized using the following categories: (1) if whether participants looked at the cue (e.g., actors' face or pointing hand), together with how long it lasted; (2) whether they looked at the targeted objects, together with the duration of the fixation and (3) whether the item was successfully found. Besides assigning the fixations to their ROIs, the duration of the interval between entering the room and localizing the target was determined. Upon inspection of the video and eye tracking recordings, it was decided to remove data of six (of the original 66) participants because of poor quality of either the video, the eye tracking or both. To facilitate interpretation of temporal intervals, time measurements were converted from frames to seconds. For most of the statistical comparisons linear mixed-effects models were used with participants as a random factor. Over the course of the experiment, it was observed that participants often did not find the target within the 1-minute search time, and as a consequence, many of the observations were censored at this interval. Such censored observations were not uniform across conditions, and therefore additional analyses were needed to deal with this aspect of the data in the form of survival analysis. The following section provides a brief introduction to survival analysis. To conduct the survival analysis and visualize the results two R packages were used: "survival" (Therneau, 2015) and "survminer" (Kassambara, Kosinski, Biecek, & Fabian, 2019).

4.3.1.7 Survival Analysis

Survival analysis (SA) - also known as duration analysis, event history analysis, failure time analysis and hazard analysis - is an ensemble of statistical methods to explore when and whether an event or events occur (P. D. Allison, 2010; Greenhouse, Stangl, & Bromberg, 1989; Keiley & Martin, 2005; Miller Jr, 2011; Panis & Hermens, 2014; Singer & Willett, 1991, 2003). Traditionally visual search results are analysed by examining the average successful search times (to locate the target) and error rates (target not found, or incorrect target selected). Such analyses rely on most trials ending in the successful localization of the target, and relatively small differences between how often the target was found across conditions. Both requirements were violated in the present experiment.

Survival analysis provides a method to jointly analyse search times and success rates, as well as incorporating trials in which the target was not found within the given time (P. D. Allison, 2010). Survival analysis constructs a model distribution of the time needed to find the target (the survival curve), and a model distribution of the conditional accuracy function (CAFs: plotting the accuracy as a function of search time speed, for reference see Stins, Polderman, Boomsma, and de Geus (2007)) the influence of search time latency to accuracy of observers responses (i.e., successfully locating an item in a room).

Here, the event is finding the target, and as a consequence the survival time was the interval between entering the room and finding the target. One of the most important feature of survival analysis is that it deals with right-censored observations (Panis & Hermens, 2014). Here, these are the trials in which participants failed to find the target item or both within the study's response interval (i.e., 1 minute) and their search time is longer than a specific value. In a traditional analysis unsuccessful trials are removed from the analysis. However, such removal of these trials might lead to the loss of valuable information about participants' behaviour. Moreover, removing trials where the target was not found, or setting the search time to the end of the allocated time (i.e., 1 minute) would lead to an artificially low average time (in fact, participants needed more than a minute, but it is unclear how much more).

Survival analysis uses two key functions: (a) the survival probability, also known as survivor function (S(t)) and (b) the hazard function or failure rate (h(t)). The survivor function represents the probability that participants has stopped surviving (by finding the targets) at time t. The hazard function represents the chance of the event at time t, given that the event has not occurred yet (i.e., the instantaneous chance of finding the target at time t, given that it has not been found yet).

For our survival data two models were used: (1) the Log-Rank test and (2) the Cox proportional hazards regression model. The Log-Rank test is a widely used method to compare two or more survival curves. As a null hypothesis the model assumes that there is no difference between the curves (e.g., the chances of finding the target at time t do not depend on the cue in the room or information about the presence of the cues). The log rank test compares the number of events in each group to what would be expected if the null hypothesis was true. The resulting statistic is distributed with a Chi-square distribution (likelihood ratio test). To perform the Log-Rank test, R's survdiff function was used (survival package). This function uses information about the number of subjects in each group (N), the weighted observed number of events in each group (Observed), the weighted expected number of events in each group (Expected) and computes a Chi-square statistic and a p-value.

The Cox proportional hazard regression model is essentially a regression model, which allows the simultaneous evaluation of the effect of several factors on survival (e.g., finding the targets in the room). In other words, Cox models explore which specific factors influence the hazard rate. The Cox proportional hazard regression function in R will provide the z-value, regression coefficients (beta coefficient), hazard ratios hazard ratios confidence intervals and global statistical significance. It will provide this for three tests (the likelihood-ratio test, score log rank statistics and the Wald test). When sample size is small the recommended test is the likelihood-ratio test.

4.3.2 Results

4.3.2.1 Cue fixations

Figure 4.9 shows the percentage of participants that looked at each of the cues at least once (left bar: percentage of participants who did not looked at the cues, right bar: percentage of participants who looked at the cues). A Chi-square test of independence showed a statistically significant difference between the percentages of looking or not the cues ($\chi^2(1) = 4.27, p = 0.04$).

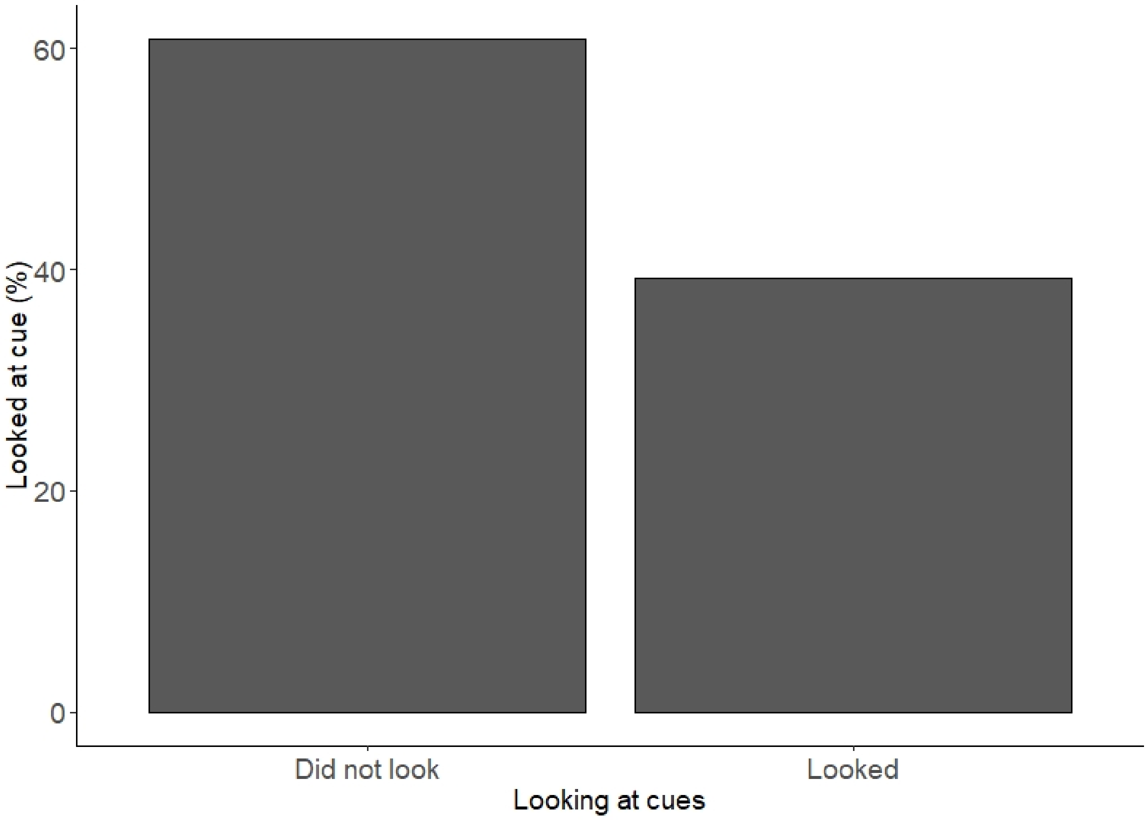

Fig. 4.9 Percentage of participants looking at least once at the cue, pooled across cues and the two groups (informed about the cues and did not informed about the cues).

Figure 4.10 shows that whether participants looked at the cues depended on the type of cue and the group participants belonged to. Graph suggest that arrows were looked less often, compared to the two social cues, but being informed about the cue influenced the difference between looking at the arrow.

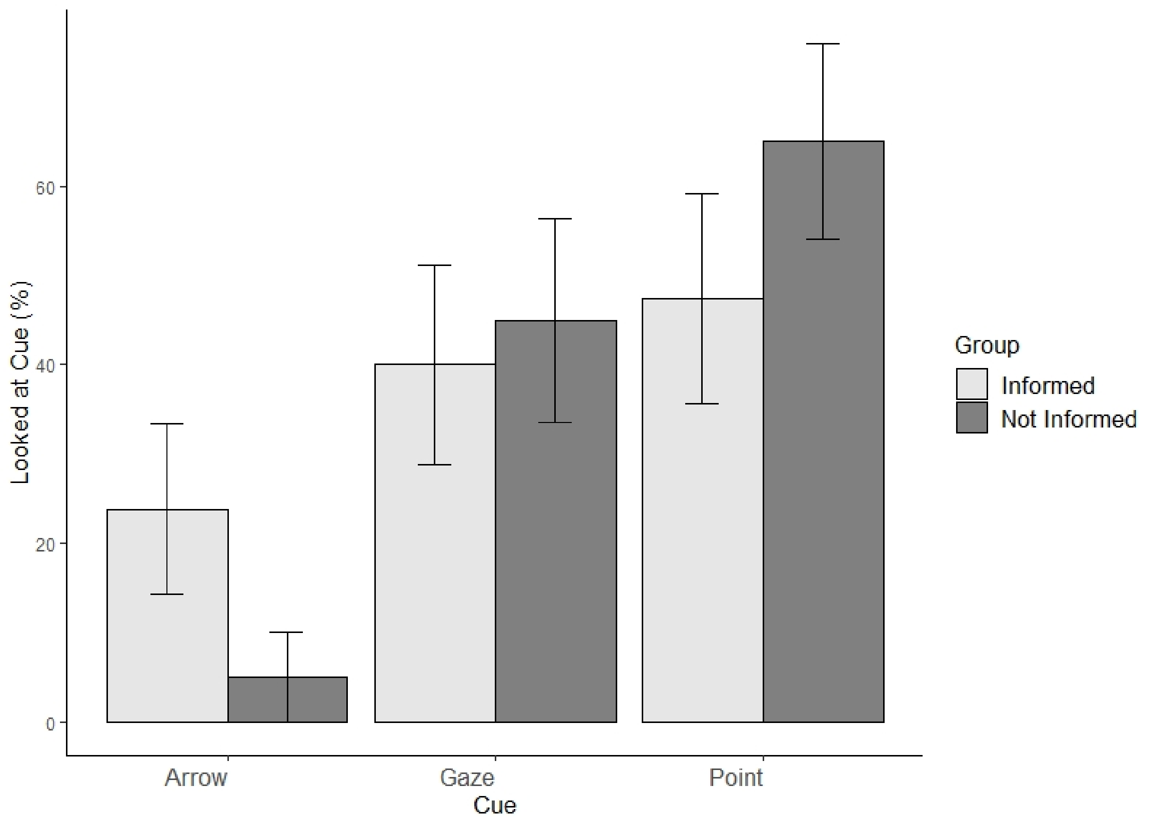

Fig. 4.10 Percentage of participants looking at the different types of cues. Left: people informed about the cues "Informed", Right: people not informed about the cues: "Not Informed"). Error bars represent the standard error of the mean across participants.

A binomial mixed effect logistic regression comparing the groups (knowing or not about the cues) for looking the cues or not and the three cues suggested no interaction between these factors ($\chi^2(5) = 2.58, p = 0.28$). Binomial mixed effect logistic regression for the main effect of the group on looking at the cues or not, showed no significant results ($\chi^2(1) = 0.04, p = 0.85$). Pairwise comparisons, using a binomial mixed effect logistic regression, on the differences between looking at the cues or not and when people were or not informed about the cues showed no significant difference for the two social cue and a significant difference for the arrows ($\chi^2(1) = 5.24, p = 0.02$). Same analysis exploring the main effect of the cues in each group showed significant results for the group were not informed about the cues ($\chi^2(2) = 14.34, p = 0.001$) but no significant results for the group where participants were informed about the cues ($\chi^2(2) = 3.41, p = 0.18$). Pairwise comparisons (mixed effect binomial logistic regression) for cue differences in the group where participants were not informed about the cues significant differences between the arrow cues and the other two social cues (arrow vs gaze: $\chi^2(1) = 6.53, p = 0.01$; arrow vs point: $\chi^2(1) = 25.65, p < 0.0001$) and no significant differences between the two social cues (gaze vs point: $\chi^2(1) = 1.63, p = 0.20$). Same analysis showed no differences between the cues when people were hinted about the cues (arrow vs gaze: $\chi^2(1) = 2.07, p = 0.15$; arrow vs point: $\chi^2(1) = 2.70, p = 0.10$; gaze vs point: $\chi^2(1) = 0.04, p = 0.85$).

4.3.2.2 Looking time on the cues

Figure 4.11 plots how long each group participants looked at the cue, when they looked at the cue (participants who did not look at the cues were excluded). The plot suggests that cues were looked at for longer when participants were informed about the cues, and that gaze cues were looked at for longer than arrow and pointing cues.

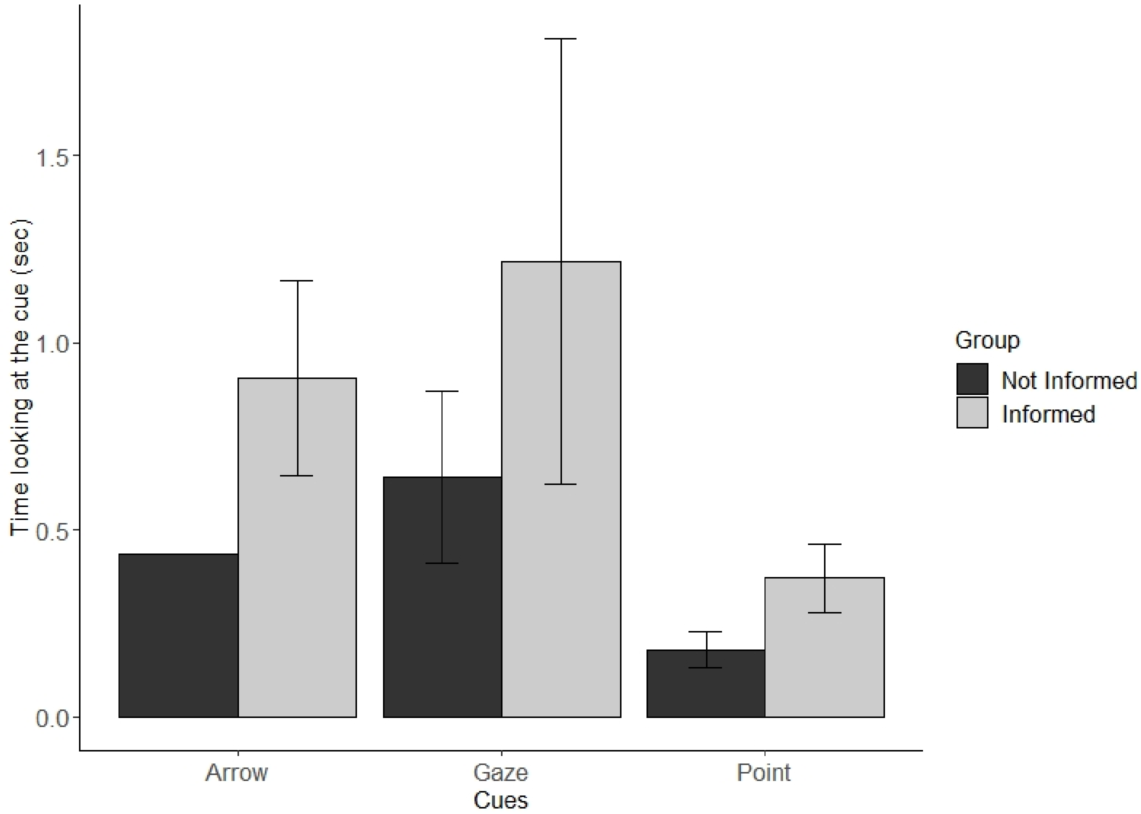

Fig. 4.11 Total looking time (sec) on the cues (aggregated across multiple looks) comparing people informed ('Informed') and those not informed ('Not Informed'), and different cues (horizontal axis). Error bars show the standard error of the mean across participants.

Looking times on the cues showed no significant interaction between the type of cue and being informed about the cues ($\chi^2(5) = 8.70, p = 0.12$). The main effect of the two groups (being informed or not about the cues) on cues' looking time was not significant ($\chi^2(1) = 3.39, p = 0.24$). No main effect of the cue was found for being informed about the cues ($\chi^2(2) = 2.76, p = 0.25$). Marginal statistically significant results were found for the main effect of the cues for the not informed about the cues group ($\chi^2(2) = 6.16, p = 0.05$). As only one participant looked the arrow cue for the "not informed group", no comparison is possible between this cue and the other two social cues in the same group as well as the arrow in the "informed group". Mixed effect comparison analysis in each group are shown in table (4.4).

Table 4.4 Paired comparisons(χ^2-value,p-value) of total looking times on the cues across the two groups (Informed, Not Informed)

Cues	Group	χ^2-value	p-value
Gaze vs Point	Not Informed	5.28	0.02
Arrow vs Gaze	Informed	0.55	0.46
Arrow vs Point		35.68	<0.001
Gaze vs Point		2.33	0.01

4.3.2.3 Outgoing saccades

Figure 4.12 examines how often a fixation of the cue was followed by a fixation on the target (saccades analysis, see Chapters 2 and 3). The plot suggests that pointing cues were much more effective at directing the participant's attention towards the target.

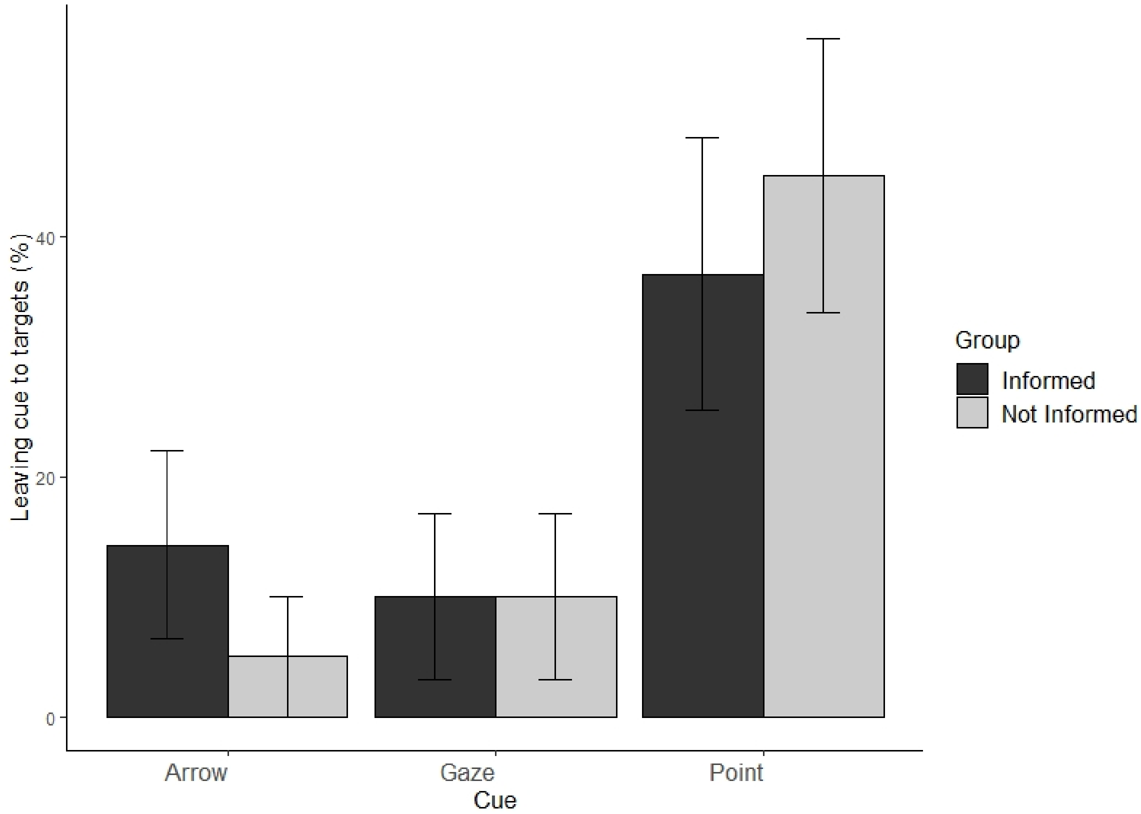

Fig. 4.12 Percentage of saccades leading to the target for the three cue types (arrow, gazing, pointing) and the two groups (informed, not informed). Error bars represent the standard error of the mean across participants.

Mixed effect logistic regression exploring the interaction between the three cues and the two groups, showed no significant results ($\chi^2(5) = 1.03, p = 0.60$).

Mixed effect analysis on the main effect of the cues on the no clue group revealed a significant difference ($\chi^2(2) = 11.64, p = 0.003$) but no significant difference for the clue group ($\chi^2(2) = 4.98, p = 0.08$). Comparing (using a binomial logistic regression) how prior knowledge about the existence of the cues influenced cues' directional effect showed no significant difference. For people not informed about the cues, there was a significant difference in the effectiveness of the pointing cues compared to the other cues (outgoing saccades point vs gaze: $\chi^2(1) = 6.53, p = 0.01$; point vs arrow: $\chi^2(1) = 19.96, p < 0.0001$; gaze vs arrow: $\chi^2(1) = 0.01, p = 0.91$. Similarly, when people informed about the cues, pointing cues were more effective than the other two cues (outgoing saccades arrow vs point: $\chi^2(1) = 14.81, p = 0.0001$; gaze vs point: $\chi^2(1) = 4.12, p = 0.04$; arrow vs gaze: $\chi^2(1) = 0.18, p = 0.67$).

4.3.2.4 Target localization

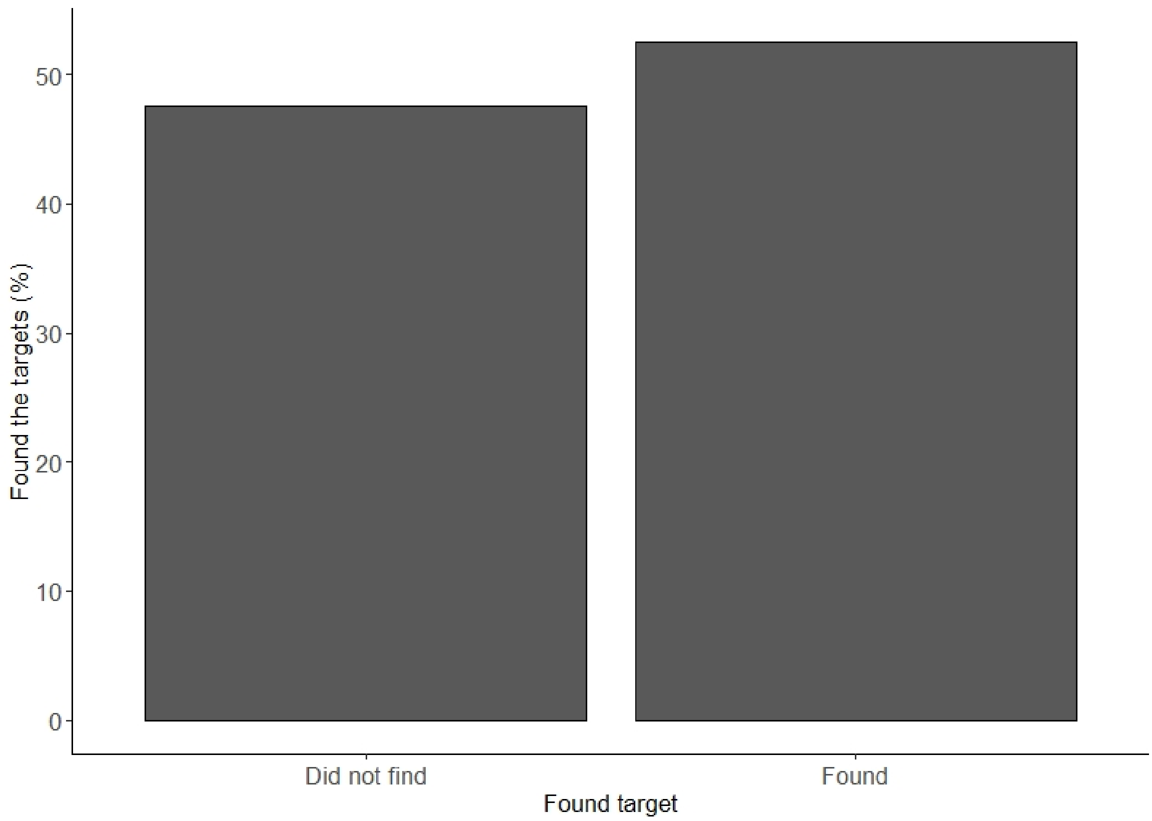

Fig. 4.13 Percentages of participants successfully locating or not the targets, pooled across targets and the two groups (informed about the cues and did not informed about the cues)

Figure 4.13 shows the percentages of success rate to locate both targets. The plot suggests that no difference between locating the target or not. This was confirmed with a Chi-square analysis ($\chi^2(1) = 0.16, p = 0.69$). Table (4.5) shows the percentages of successful target localization for each target separately. The table suggests slightly better success to locate the second target. However, Chi-square analysis on the percentages of successful target localization showed no difference between the two targets ($\chi^2(1) = 0.45, p = 0.51$). Similar analysis showed not difference between the two targets for the percentages of failing to locate the targets ($\chi^2(1) = 0.49, p = 0.48$). No difference was found between the percentages of successfully locating or not the target for each target (Target 1: $\chi^2(1) = 0.11, p = 0.74$; Target 2: $\chi^2(1) = 1.08, p = 0.30$)

Table 4.5 Percentages of participants successfully locating or not the two targets in detail.

Target	Found	Not found
Target 1	46.67%	53.33%
Target 2	58.33%	41.67%

Figure 4.14 shows how often participants (both groups) found the targets and under each cue. The plot suggests that hinting about the cues aided the localization of the targets (higher percentages of the 'Informed' group). The plot also suggests that under the pointing cue, targets were located more successfully. Note here that graph and the following analyses include both targets as detailed analyses (comparing the two targets) was not permitted. Breaking down the data to conduct a more detailed analyses created a huge inequality of the sample size per condition, which neither the linear mixed-effects models nor the traditional ANOVA could cope with. Ignoring this restriction, would have resulted to reach to wrong outcome and conclusion for participants' behaviour and cues effects.

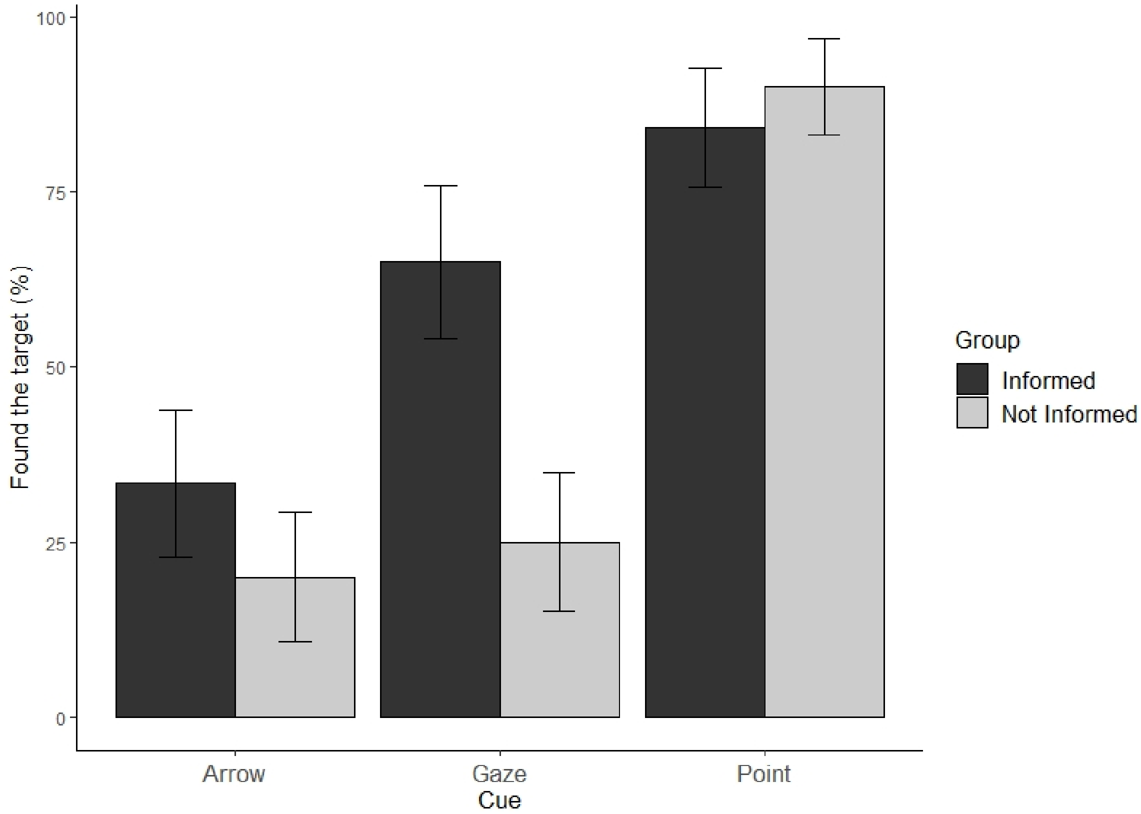

Fig. 4.14 Percentage of success targets' localization when informed ('Informed') or not informed ('Not Informed') about the cues per cue conditions. Error bars represent the standard error of the mean across participants.

There was a significant interaction between being informed about the cues and the three cues for the success rates on targets' localization (mixed effects logistic regression: $\chi^2(5) = 72.07, p < 0.0001$). The effect of the cues in each group (using a mixed effect logistic regression) for the success rates on targets' localization, showed a significant difference for both groups (Informed: $\chi^2(2) = 30.39, p < 0.0001$; Not Informed: $\chi^2(2) = 40.10, p < 0.0001$). Mixed effect logistic regression analysis for the effect of the groups on cueing condition, showed no significant differences for the arrows and pointing cues (arrow: ($\chi^2(1) = 0.94, p = 0.33$; point: $\chi^2(1) = 0.1, p = 0.91$), but significant difference for gazing ($\chi^2(1) = 6.66, p = 0.01$). Differences between the three cues for the "Informed" group were also computed. Mixed effect logistic regression showed significant differences between the arrow and the two social cues (arrow vs gaze: $\chi^2(1) = 14.62, p = 0.0001$; arrow vs point: $\chi^2(1) = 24.25, p < 0.0001$) and no significant differences between the two social cues (gaze vs point: $\chi^2(1) = 2.18, p = 0.14$). Same analyses were conducted to explore the differences between cues in the "Not Informed" group, showing significant differences between the pointing cues and the other two cues (arrow vs point: $\chi^2(1) = 36.57, p < 0.0001$; gaze vs point: $\chi^2(1) = 33.07, p < 0.0001$) but no differences between the arrow and gaze cues ($\chi^2(1) = 0.01, p = 0.93$).

4.3.2.5 Time to locate the targets

Figure 4.15 shows the survival curves showing the time until finding the two targets. Note that these curves do not end at zero, but at a different value, indicating the proportion of trials in which the targets were not found.

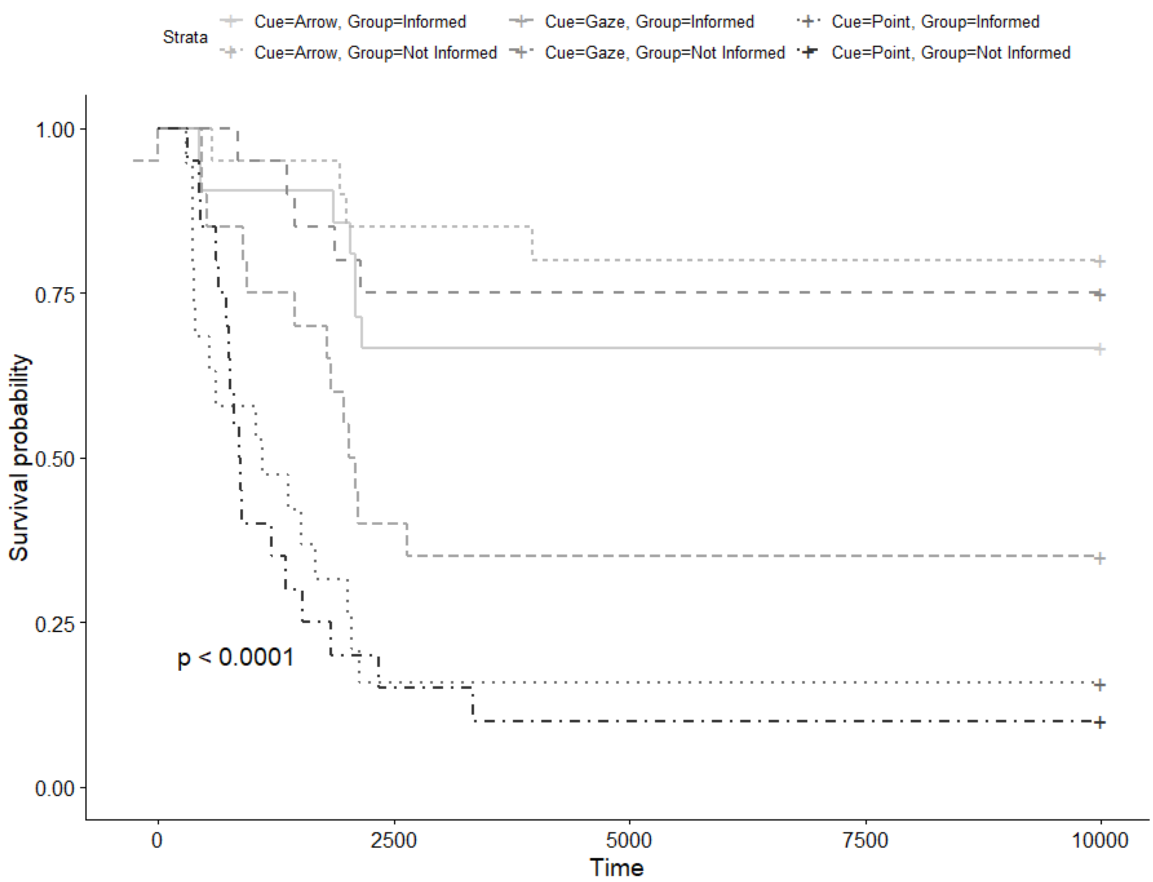

Fig. 4.15 Survival curves showing how long participants took to locate the targets, and the proportion of trials where the targets were not found (asymptotic values). Lines show the different cue and group combinations (arrow, gaze, point; and "Informed", "Not Informed").

In order to assess if this observation is reliable (statistically significant), a log-rank test showed that survival times differed across the combinations of cues and information about the cues ($\chi^2(5) = 54.1, p < 0.0001$; for more information on the analysis, see Table (4.6)).

Table 4.6 Log-rank test values for comparisons of time localization for the three cues across the two groups (Informed, Not Informed)

Cues & Group	N	Observed	Expected	$\frac{(O-E)^2}{E}$	$\frac{(O-E)^2}{V}$
Arrow, "Informed"	21	7	13.38	3.04	3.88
Arrow, "Not Informed"	20	4	13.71	6.88	8.85
Gaze, "Informed"	20	13	10.13	0.81	0.97
Gaze, "Not Informed"	20	5	13.14	5.04	6.40
Point, "Informed"	19	16	6.23	15.32	17.17
Point, "Not Informed"	20	18	6.41	20.98	23.92

Because of a rather small number of observations (see Table (4.6)), the analysis was repeated with the Cox proportional hazards model. In line with above results, Figure 4.16 suggests that pointing cues led participants to more often find the targets but knowing about the presence of the cues did not.

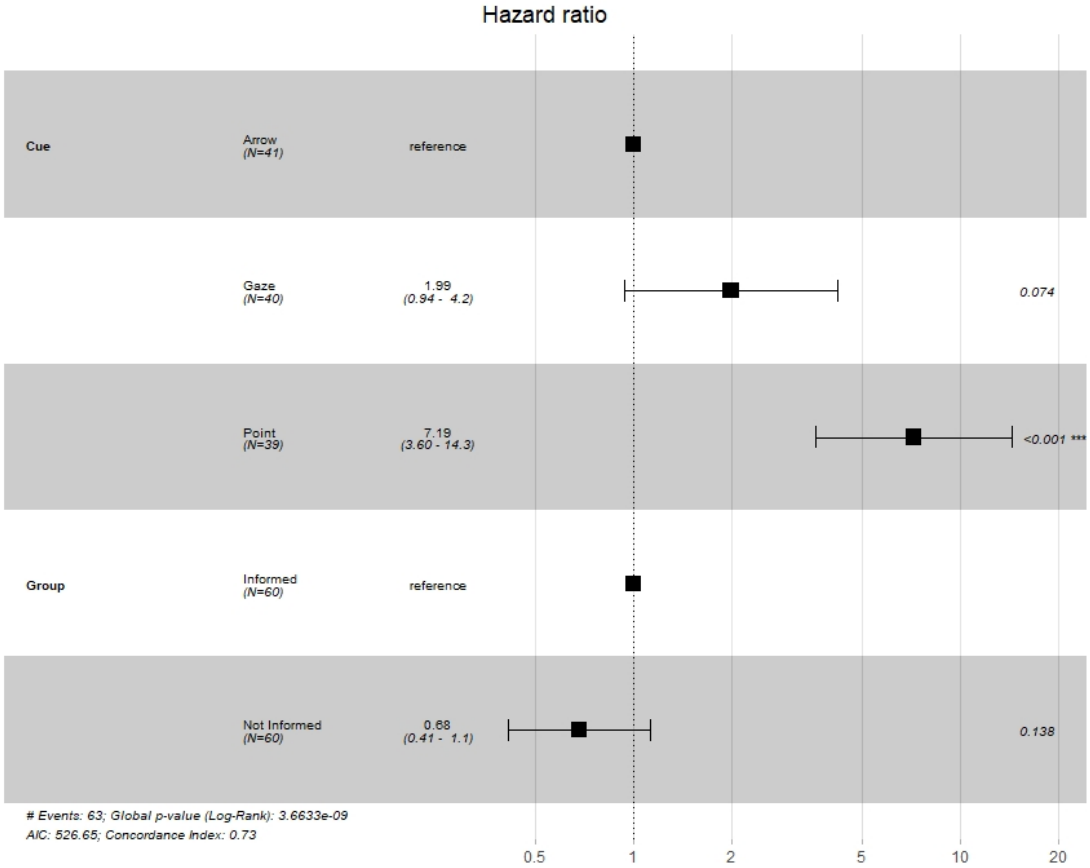

Fig. 4.16 Forest plot output for interaction of locating the targets for the three cues and two groups. Forest plot for the effect of the three cues and two groups on the hazard ratio of targets localization (Fixed-effects model)

Table (4.7) presents the results from the Cox proportional hazards model, using the arrow and cue informed group as references. Results from the analysis showed that under the pointing cue, targets were found faster. Knowing about the existence of the cues did not help participants locate the targets faster. Similar findings were obtained when gazing cue was used as a reference and compared with pointing ($z = 4.36,\ p > 0.0001,\ beta(coef) = 1.27,\ Hazard\ ratio = 0.28,\ SE(coef) = 0.30$)

Table 4.7 Statistics from Cox proportional hazards model (beta coefficient, hazard ratio, standard error, z-values and p-value) for the interaction between cues and groups for success rate in target localization.

Comparison	beta-coef	Hazard ratio	SE(coef)	z-value	p-value
Gaze	0.69	1.99	0.38	1.79	0.07
Point	1.97	7.19	0.35	5.60	>0.0001
No Clue	-0.38	0.68	0.26	-1.49	0.14

*Likelihood ratio test = 42.19 on 3 $df, p > 0.0001$

4.3.2.6 Target total looking times

Figure 4.17 shows how long targets were looked at (if they were fixated), suggesting that participants looked at the first target for longer.

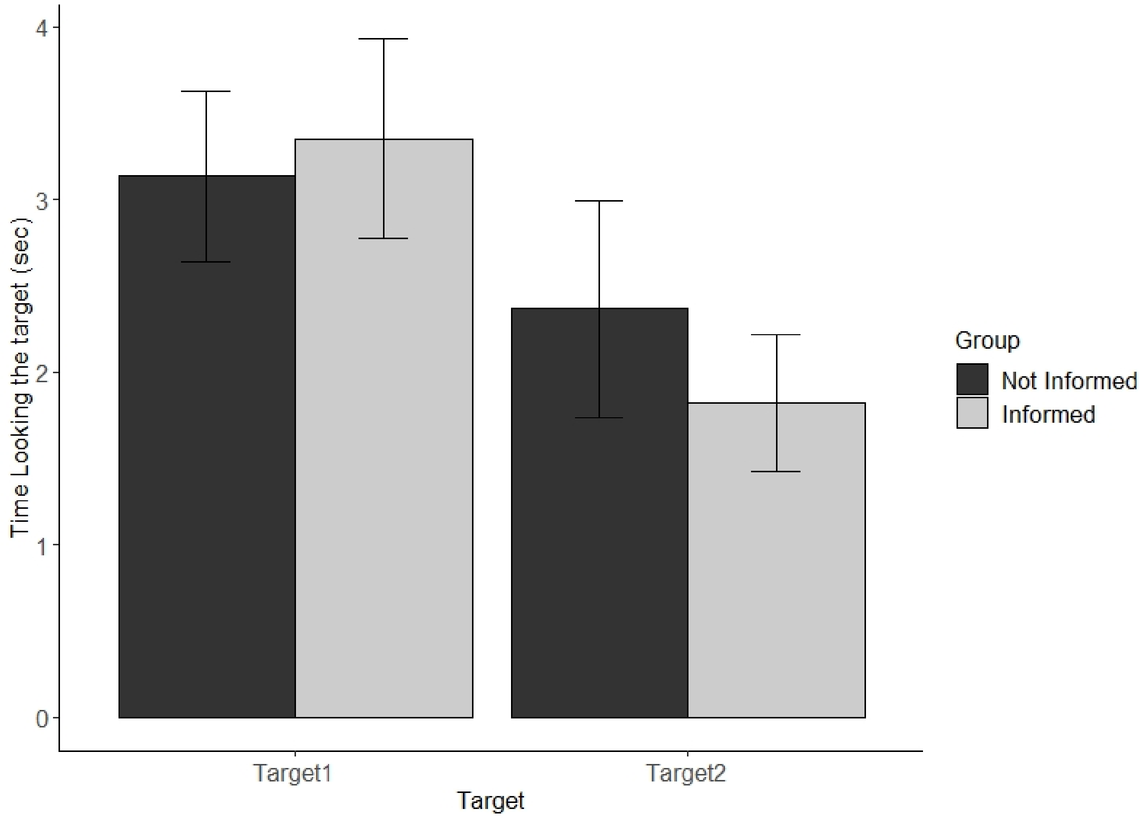

Fig. 4.17 Total looking time at the two targets and groups differing in being informed about the cues, pooled across the cueing conditions (arrow, point and gazing). Error bars represent the standard error of the mean across participants.

There was a marginal significant interaction was found between being informed about the cues and the number of the target on target looking time ($\chi^2(3) = 8.00, p = 0.05$). A significant effect of target number was found ($\chi^2(1) = 7.68, p = 0.01$), but no effect of being informed about the cues ($\chi^2(1) = 0.19, p = 0.66$). Comparison analysis between the looking time on the targets in each target for the "Informed" group showed a significant difference ($\chi^2(1) = 7.10, p = 0.01$) but no significant difference between the two targets for the "Not Informed" group ($\chi^2(1) = 0.90, p = 0.34$).

4.3.3 Discussion

Studies comparing the effects of social (gaze and pointing) and symbolic cues (arrows) on observers' attention have typically employed a computer-based task showing static cues in isolation and at fixation. Studies examining whether people look at other people in real-world tasks suggest that while observers look at people on a computer screen, they tend to avoid looking at people in the real-world (Gallup, Chong, & Couzin, 2012; Gallup, Hale, et al., 2012; Laidlaw et al., 2011). Social attention experiments that make use of the standard cueing paradigm may therefore provide limited information about how social attention works in the real-world. The present two experiments provided the first attempt to compare social and symbolic cues in a real-world setting.

Experiment 5 was a first attempt to compare social and symbolic cues in real-world settings. Cues were presented as drawings, printed on a sheet of paper and placed in four different rooms. Participants entered rooms being presented a search target in a photograph, where cues were placed to point at the target. In such a context, social cues were found to be looked at more frequently than arrows. However, cues were often ignored and did not substantially speed up the localization of the target. A possible reason is that targets were relatively easy to find: Search times were a matter of seconds, and targets were always found. Targets were located at expected locations (e.g., a cup on the counter), not hidden away from view, and the rooms were not very cluttered. Previous studies (e.g., Mack & Eckstein, 2011) exploring visual search in natural environments have shown the same effects. Similar to the findings of this experiment, these studies have found that targeted items were located faster when they were encountered in an expected location compared to an unexpected location.

Another concern was that cues were lacking ecological validity, being line drawings on a sheet of paper. Particularly, gaze cues were occasionally presented vertically and lacked any additional characteristic which could be recognized as eyes (e.g., eye lashes, eye brows). This might have looked peculiar to the participants and participants might have not recognized these lines of drawing as eyes. However, previous lab based studies (e.g., Bayliss & Tipper, 2006a) presented gaze cues in a similar manner (e.g., vertical eyes without any other facial characteristic) and have found that even in this way gaze cues can show a strong cueing effect. A further possible issue was that the control condition did not include a sheet of paper. Search times were in fact slightly shorter for the cues compared to the no cue condition, and the mere presence of a sheet of paper next to the target may have helped. Better ecological validity would have been obtained with actual people rather than print-outs of hands and eyes to provide the cues. In addition to the experimenter, two further people would need to be present each time a participant is tested. Having actual people inside each room may also prompt surprise or confusion in the participants, as they may be uncertain why these people are in the room.

Experiment 6 tried to address these possible issues by using actual people as cues and make it more difficult to find the targets by not providing an image of the target, but merely the naming it. Moreover, actors were actively part of the experiment to avoid any unexpected effect from suddenly seeing a person in the room. Interestingly, cues were still not looked at often, even when participants were aware of cues' existence while searching. These findings are in line with past studies (Gallup, Chong, & Couzin, 2012; Gallup, Hale, et al., 2012; Laidlaw et al., 2011) suggesting that people avoid looking at other people in a real-world setting. Because experiment 5 showed that people also did not look at cues at sheets of papers, a more plausible reason may be that people were too engaged in the search task to look at the cues. Moreover, as both experiments took part in real-world settings it is quite possible that cues were indeed noticed but through participants' peripheral vision. However, this is merely a speculation as mobile eye trackers are not keen to measure this part of vision.

One indication is from participants self-reports were the majority in both experiments reported that they were aware of cues' existence. Feedback from the participants after performing the task suggests that another factor may have played a role: They often reported that they thought the actors in the room were there to confuse them or "trick" them. This may be related to the sample. The majority of participants were psychology students, who have experience in taking part in studies that aim to test something not revealed or purposely hidden at the start of the experiment. Previous studies (e.g., Freeth et al., 2013) have suggested that people spend more time looking at another person's eyes and face, when these regions convey information essential to the understanding of a social situation (e.g., asking a question). To see if this will be the case in the current experiment, people in the room were given a role (i.e., help them locate the targeted item). Interestingly, this impression was not altered by informing participants that the other people in the room were there to help them find the targets. While the room was relatively cluttered with items, the actors were the only other people in the room. It would therefore be interesting if pointing or gazing people would be more effective if there were other people not providing such cues as well (e.g., a restaurant or airport setting).

Another possible reason for the lack of a significant difference is a possible lack of statistical power. Higher statistical power could have allowed us to detect any small differences between cues that may be present. Participants were exposed to a limited number of trials (e.g., entering a room only once or twice), and a limited number of participants could be tested to compensate for the increased within-subject variability that was the consequence of this small number of trials per participant. In lab based experiments, a larger number of trials can easily be presented, as it simply involves more images or videos. However, in real-world experiments increasing the number of trials is not always feasible (e.g.,limited number of rooms available). It should be noted, however, that for a real-world experiment, the number of participants tested in Experiment 5 and 6, was already quite high. For example previous real-world studies (MacDonald & Tatler, 2013), exploring the role of gaze cues in collaboration, had only twenty-four participants for a single trial. Even though the number of participant in the current experiments was high, still increasing the number of participants per group might have been needed to compensate for any issues resulted from the data loss caused by the use of the mobile eye tracker and the experimental design.

As well as exploring the time spent looking at the cues, the present studies also explored the relationship between cues and success in target localization. In Experiment 5 prior knowledge of items' characteristics were provided. Therefore, it was not surprising to see, participants always finding the item without any help from the cues. However, Experiment 6, where both targeted items were unknown and the existence about the cues were manipulated, had lower success ratings (around 25% for first target and 29% for second target). This seems to be independent of whether participants were informed about the existence of a help to locate the items, suggesting that hinting or not about the existence of the cues did not help to overcome the lack of knowledge about the item they were looking for. In both groups and under the pointing cue, targets were found faster and more successfully compared to the other two cues. This is not surprising as pointing has a more distinct movement whereas gazing can be ignored, not noticed and can mislead participants to identify the right item. This can also be reflected from the outgoing saccades analyses were pointing led to more successful saccades than any of the other cues. The advantages of the pointing cues to successfully direct the observers' attention to a target has also been found in the previous chapters in this thesis (Chapter 2, Experiment 2 3; Chapter 3). Arrows failure to lead faster to the targets can be related to the position it was placed in the room. Arrows were placed directly above both targets and it was expected to help participants more in their search. Potentially the arrows were less salient from the other two cues which had a clearer shape and were easier to notice. To enhance arrows' presence arrow were place in a white background, it was of bright colour and at the level of eye sight. However, this did not help, and arrows were noticed less and showed no effect to increase targets' localization. From the participants who noticed the arrow, all of them reported that arrows were considered as part of the environment and not the experiment. Although this merely a speculation, it can be an indication which can help in the design of a future study.

Interestingly, Experiment 6 found limited cueing effects from the gaze cues. This is in stark contrast to more traditional gaze cueing studies that found strong effect of gaze cues (e.g.,Driver et al., 1999; Friesen & Kingstone, 1998). A possible reason for not finding the gaze cueing effect here is that cues were embedded in natural scenes. Static gaze cues in such scenes were not particularly effective (e.g.,Hermens & Walker, 2015; Chapter 2 and 3). This cueing effect of gaze cues did not improve when people rather than printed cues were used, suggesting that gaze cues, while effective in isolation, are not effective in a cluttered and naturalistic environment. A possible reason why gaze cues are less effective in a real-world setting is that it is often unclear what object the person is looking at. Generally, people are fairly poor at estimating the direction of a cue in a 3D image (e.g.,Doumen et al., 2010). Gaze cues are different from arrow and pointing cues in that the actor does not typically gaze at a target object in order to direct the observer to that object. If someone wants to point someone else to an object, they will typically use a pointing gesture. Gaze cues therefore mostly indicate an interest in the person looking without the direct intention to communicate that interest. The lack of such clear intention may reduce the effect of the gaze cue particularly in a search task, where the observer may only pay attention to cues that clearly present a hint towards the target.

4.4 Conclusion

This chapter was one of the first to explore social attention in a real-world setting, where the observer was embedded in a scene containing the social (and symbolic) cues. When observers are given a search task, they pay little attention to the cues in the scene and do not seem to make use of the cues for finding the target, even if the purpose of the cues is made clear to them. The experiments presented here show that exploring social attention in the real-world is possible, and future studies may adopt a similar approach to study whether social cues in a real-world setting have stronger effects if observers have a less defined task at hand.

Chapter 5

General Discussion

Human eyes are considered one of the most important body parts, employed in the service of communication. Eyes are the object of interest from the early age of 2 months old (Maurer, 1985) and even from the early age of 12 month old, infants are keen to follow another's person gaze (Thoermer & Sodian, 2001). A plethora of studies (e.g., Driver et al., 1999; Friesen & Kingstone, 1998; Langton & Bruce, 1999) have explored the effects of gazing cues to human attention, postulating a robust cueing effect even when eyes' direction is uninformative to a certain target's location. However, there are few studies which have suggested that the directional effect gazing cues have is not unique, and other social (e.g., pointing hands) and non-social cues (e.g., arrows) can show similar effects (e.g.,Burton et al., 2009; B. S. Gibson & Bryant, 2005; B. S. Gibson & Kingstone, 2006; Langton & Bruce, 2000).

More recently a debate has begun, discussing how the methodology (e.g., empty background, cropped heads and hands) used in these studies can reflect the effects social and symbolic cues have in the real life. Very few studies have attempted to explore the effects of social and symbolic cues when they are part of natural static scenes (e.g., Birmingham et al., 2009; Hermens & Walker, 2015) or in the real-world (e.g.,Gallup, Chong, & Couzin, 2012; Laidlaw et al., 2011). However, the number of these studies is limited and their results inconclusive. Therefore it is still unknown how the difference between the two methodologies (ecologically valid and non -ecologically valid) can influence differently observer's attention towards the cues. In addition it is still unclear if under more realistic environments both social and symbolic cues will affect the observers' attention differently compared to the classic empty environments used in previous studies.

This thesis aimed to explore these differences and further our understanding on how social and symbolic cues influence the observers' attention when embedded in natural scenes and the real-world. Particularly this thesis aimed to explore 1) if gaze cues can still capture people's attention and 2) elicit a strong cueing effect when presented under a naturalistic environment and compared with another social (pointing hands) and symbolic cue (arrows). In order to explore these two research questions, actual people, rather than cropped hands and faces, were used to provide the two social cues (gazing and pointing). All cues were embedded in natural scene environments (studies in Chapter 2 and 3) and the real-world (studies in Chapter 4). The effects of social and symbolic cues were also explored under different scenarios (multiple people, search tasks) to investigate whether under these conditions, cues can influence people's attention differently (e.g., time spend looking at them; cueing effect). In addition, all the studies in this thesis attempted to explore social attention following an ecological valid approach to identify important methodological limitations, which potentially can improve the way social and symbolic cues will be investigated in the future.

5.1 Summary of studies

Studies in Chapter 2 aimed to investigate the effects of social and symbolic cues when embedded in static natural scenes. Throughout the three experiments, cues were all presented in a static form. Moreover, both social and symbolic cues were presented under three different scenarios, a single cue under a free viewing task (see Experiment 1); multiple cues pointing at the same target or one pointing away from the target (see Experiment 2) and finally a visual search task to explore whether task effect can influence eye movements towards the cues differently (e.g., less time spend on the gaze cues). Results in all three studies showed that gaze social cues consistently strongly captured people's attention compared to a symbolic cue, like an arrow. This finding was in agreement with the previous literature (e.g.,Hermens & Walker, 2015; Zwickel & Võ, 2010) using similar stimuli (e.g., natural scenes, people producing the social cues). When presented with another social cue, like pointing, gaze cues showed less effect. Potentially this is related to the fact that when presented together, pointing and gazing cues are equally important for the observer, to decipher the information provided by the cues (e.g.,Langton & Bruce, 2000).

In terms of the cueing effect, this strong attentional shift related to gazing cues previous studies reported (e.g.,Driver et al., 1999; Friesen & Kingstone, 1998), was not found. Instead, other cues indicating direction to a certain location or item (e.g., arrows and pointing hand) showed to strongly shift the observers' attention. This is not surprising considering that in real life encounters, arrows and people pointing, could be considered as indicators of direction towards locations or targets. On the other hand, gazing cues communicate more than one message (i.e., a person's thoughts, emotions) apart from direction (e.g.,Baron-Cohen, 1995) and when encountered in real life they will be usually accompanied by actors' pointing hands and some verbal instructions (i.e., think when asking directions to locate the boarding gate in an airport).

Chapter 3 moved a step closer to present cues as the observer will encounter them in real life situations. In day to day life when encountering social cues, it is most likely to see the entire action (hand and body moving) rather than the result of this movement, which static cues represented. Therefore, a study in Chapter 3 presented dynamic social and symbolic cues under two different tasks (free view and memory) to explore if motion and task will affect observers' attention differently towards the cues. Similar to the previous studies in Chapter 2, gaze cues attracted most of observers' attention but only in the absence of motion. While in motion, pointing was looked at the most, suggesting that motion might be a factor influencing attention towards the cues. In agreement with previous natural scenes' studies (e.g.,Hermens & Walker, 2015), task did not influence the time the observers' spent looking at the cues. Similar to Chapter 2, gaze cues failed to shift attention with pointing (a cue with more distinct move and a meaning associated with direction) being the one successfully directing participants' attention to a certain target. In contrast to the previous literature (e.g,Fletcher-Watson et al., 2008; Itier et al., 2007), shifting of attention was not influenced by the tasks, potentially indicating that the task effect was overcome by cues motions. The way arrows were presented was a profound limitation on this study which will be discussed in the following section (see 5.2).

The final two studies in Chapter 4 explored social attention in a real-world environment. In the previous two chapters participants were seated with their head fixed and the cues were presented in a computer monitor. However, in real life when people are looking at these cues, they have the opportunity to adjust their position in order to decide where and what these cues are pointing at. Moreover, both social and symbolic cues can be encountered under different circumstances: (1) provided by actual humans or (2) signs. The last two experiments aimed to explore a real life situation where people are searching for familiar and unfamiliar items. Two different versions of social and symbolic cues (cues drawn on a piece of paper and social cues provided by actual people) were used. Although stimuli in the two studies were different, findings were consistent with the previous literature (e.g.,Gallup, Chong, & Couzin, 2012) where it was found that gaze cues were ignored. In fact, both studies revealed that while searching, cues are generally ignored, suggesting that task was more important and bias people's attention towards the cues. In agreement with the previous chapters in this thesis, gazing cues did not direct or help the observers to locate the "search for" items. Instead, pointing was more successful and significantly helped to find the requested objects faster and accurately. Again, arrows did not show a significant effect, suggesting that these cues have a minimal effect in affecting people's attention and potentially are sensitive by their position presented in the environment.

5.2 Are the effects of gaze cues unique?

Previous studies on social attention (e.g.,Driver et al., 1999; Friesen & Kingstone, 1998) have strongly suggested that gaze cues capture observers' attention and show a strong cueing effect. However, all these studies have presented the cues in an empty background with nothing else to see and used cues in a way that is rarely seen in the real-world. Although findings from these studies are important to understand social attention, their lack of ecological validity approach is concerning. Therefore, the use of a more simplistic and less realistic methodology might have led to “misleading” results about cueing effects. One approach more recent studies and this thesis (e.g.,Birmingham et al., 2009; Hermens & Walker, 2015; Zwickel & Võ, 2010) followed was to explore the effects of gaze cues when these are embedded in natural scenes and the real-world. In addition, some of these studies have compared gaze cues with other social cues (pointing) and symbolic cues to determine if they will show similar or stronger effects from gazing. This thesis followed this line of work and extended the line of research.

5.2.1 Gaze cues and their ability to capture observers’ attention

One of the first questions this thesis addressed, was when cues are embedded in naturalistic environments and the social cues are produced by actual people with their body in full view, will be the one which strongly capture people’s attention. The overall results from the first two chapters suggest that people do tend to look at other people when they are part of an image either when they are gazing or pointing at a certain location. This is in agreement with previous literature (e.g.,Birmingham et al., 2008a, 2009; Fletcher-Watson et al., 2008, 2009; Hermens & Walker, 2015) which have also suggested that in the presence of people in a scene, humans show a strong tendency to look at them. Therefore it can be assumed that when it comes to using more naturalistic scene stimuli, humans are still the object of interest. This is not surprising if we consider that attending at other human’s action is an essential element of evolution and cognition.

Gaze cues showed a stronger effect to capture observers’ attention when compared with an arrow. This finding is widely supported by the literature (e.g.,Birmingham et al., 2009). A possible reason for the arrows to attract less attention is related to the characteristic of the cue (e.g., size: usually arrows are smaller than humans; their location in an environment) and the meaning these cues have been associated with. Arrows are simple as a design with their meaning only suggesting direction and location. Therefore after participants determined the direction the arrows were pointing, there was no need for them to spend more time looking at them. On the other hand, gaze cues when are produced by actual humans are more complex. Apart from indicating direction, gaze and head orientation reflects a person's feelings, intentions and other essential for communication elements (e.g.,Argyle & Cook, 1976; Baron-Cohen, 1995; Ellsworth, 1975; Kleinke, 1986). Therefore, it is not surprising that dwell times were longer on gaze cues, while participants were potentially trying to understand the message this person was conveying to them.

While gaze cues' ability to capture the observers' attention is widely supported by the literature (e.g.,Birmingham et al., 2008a, 2009), this ability seems to be merely restricted when compared with a symbolic cue, such as an arrow. This thesis revealed that in several circumstances (e.g., multiple cues, task, and motion) pointing can be a worthy "opponent" for gazing cues effect. When presented together, both social cues attracted participants' attention. This is in line with previous findings (Langton & Bruce, 2000) and is related to the fact, that when presented together, both pointing and gazing cues are equally important for the observers to retrieve information about the message actors' providing the cues are communicating. Although in these studies social cues were produced by the same actor, this thesis showed that even in the presence of two people producing the cues separately, these effects (dwell times) persist.

When cues were moving, time spent on them showed two interesting outcomes. While in motion, the dynamic pointing cue showed the longest dwell times suggesting that the observers looked at them more. In the absence of motion, gaze cues were looked the most when compared to pointing cues. As no previous studies have directly compared dynamic gazing and pointing cues, only speculations can be offered to explain these results. Movement of pointing is quite distinct and requires more time to reach its final position. On the other hand, head movement is faster and less distinct. Therefore, it may be expected to see that when dynamic, pointing was the one which captured the observers' attention. In the absence of motion, the observers lost their interest on pointing cues and gazing cues were looked at more. This indicates that when seeing the whole process of cueing rather than the outcome of it (e.g., as in static scenes) motion can potentially bias attention. The effect of arrows while in the motion stage can be merely an artefact from this study's design and will be discussed in the following section (5.3).

In the real-world, results from the last two studies of this thesis, were consistent with the previous literature (e.g.,Gallup, Chong, & Couzin, 2012), with gaze cues being ignored. In fact, as was discussed in the previous section, under an actual search task, cues played a smaller role in people's performance. These results suggest that the specific task at hand strongly influenced participants' gaze behaviour. Previous studies (e.g.,MacDonald & Tatler, 2013) in which a task was performed under a real-world scenario (e.g., cooking) have shown similar behaviour, with gaze cues, provided as a useful help, generally being ignored. This consistent behaviour between different real-world paradigms, raises an important questions: if people rarely look at the cues, do they still use the cues? There are three answers to this question. Firstly, when cues are fixated, fixations duration do not necessarily have to be long for cues to be informative (e.g.,MacDonald & Tatler, 2013). Even briefly looking at the cues can be informative enough for the people to locate the targeted item. Secondly, people do not generally look for the cues when they are performing a task (e.g.,Nummenmaa & Hietanen, 2009). Instead cues might be used in case of need (i.e., not sure about item) and for clarification reasons (e.g., to determine the position of the items in the room). Finally, cues could have been observed in peripheral vision and not by directly looking at the cues. However, due to restrictions in current technology this is merely a speculation.

A consistent finding with the previous studies in this thesis, is that in the presence of a human providing the social cues, attention to symbolic cues like arrows is limited. Again it can be argued, that arrows do not necessarily need to be looked for a long period of time for the observers to distinguish their characteristics and understand their meaning. Therefore, the meaning the arrows communicate is more direct and potentially requires less time to be comprehended. However, this does not necessarily guarantee higher rates of successful target localization, as it seems that even arrows can be ambiguous in terms of what they are pointing at. This potentially is related to the position in which the experimenter places them in the scene as well as other factors (e.g., distance from the target, colour of the background).

5.2.2 Are gaze cues strong enough to shift observers' attention?

The second question this thesis wanted to explore is whether gaze cues will show a strong cueing effect. Previous literature using more traditional methodology (e.g.,Driver et al., 1999; Friesen & Kingstone, 1998) have strongly suggested that in the presence of a gazing eyes observers follow the eyes' direction. This finding is consistent across both static and dynamic gaze cues (e.g.,Bayliss & Tipper, 2006b; Hermens & Walker, 2012; Kuhn & Tipples, 2010; Ricciardelli et al., 2002; Rutherford & Krysko, 2008; Swettenham et al., 2003). However, studies (both lab- and real-world based) in this thesis could not confirm the strong cueing by gaze cues. In fact, gaze cues' effect to shift the observers' attention was to a lesser extent.

One possible explanation is related to the environment in which gaze cues were presented. When embedded in natural scenes, where gaze cues are pointing at might be ambiguous (e.g., leading to a different item in the scene) and seems to highly depend on the complexity of the natural scene environment. Moreover, the presence of an actual human producing the gaze cue potentially communicates a more complex meaning than just directing attention to a certain location (e.g., emotions). This may explain why participants spent extensive time looking at these cues or revisit them to decipher their meaning and understand where these cues are actually looking at. Studies applying the classic cueing paradigm with isolated cues presented at fixation (e.g.,Friesen & Kingstone, 1998; Hermens & Walker, 2012) may have only observed fairly artificial gaze behaviour not usually observed in day to day life. In real life encounters, a gaze cue is usually embedded in a vast array of items where the absence of any other information (e.g., verbal instructions) can be confusing. Therefore, it seems that the strong gaze cueing observed in classic gaze cueing studies may be restricted to cues in isolation and at fixation. The real-world studies in this thesis, are in line with previous real-world studies (e.g.,Gallup, Chong, & Couzin, 2012) which showed that gaze cues are ignored and do not show a strong shifting effect. This lack of cueing effect from gaze cues occurred even when these cues could help the participant to locate an object. One may speculate that this lack of an effect may be due to how easy it was to find the target. However, this was not the case as participants hardly looked at the cues even when searching for the targeted item was more difficult.

Another reason to believe that the lack of an effect of the gaze cues was due to the particular nature of these cues, is that stronger cueing effects were found for the other cues: The arrow and the pointing cues. Particularly the pointing cue showed consistent cueing effects observed in almost all settings. The effects of arrows was less consistent and depended on the way they are presented in the scene (e.g., their position in relation to the items). These strong effects of the pointing cue and arrow may have several explanations. Firstly the shape of both cues is more distinct than that of the gaze cue, which particularly helps when the cue is perceived in peripheral vision (e.g.,Hermens, 2017; Langton & Bruce, 2000). As a result, arrows and pointing cues shift observers' attention more strongly than gaze cues. Potentially this is what is observed here.

Secondly the stronger cueing effects for arrows and pointing observed here, may be related to the meaning humans have associated these cues with (i.e., indicating direction). Whereas another person could gaze in a direction for other reasons, pointing gestures and arrows serve to direct attention in the observer to a location. Pointing cues are special in this respect, compared to arrows, because they involve an active movement of a person. An arrow could be placed in a room and no longer be relevant to the observers entering at a later stage. For example, an arrow on a fire exit sign may only become relevant when there is a fire. In comparison to gaze cues, arrows and pointing gestures provide stronger messages, but they also may be less ambiguous (i.e., it may be easier to determine exactly which object or point in space they are pointing to) and as discussed above they convey more than one message. However, it is important to note that even for arrows and pointing gestures, not many outgoing saccades led directly to the target.

Compared to the traditional gaze cueing paradigm, cues in natural scenes do not often lead to successive fixations on the target (see also, Hermens & Walker, 2015). Reasons for this could be that in the classic cueing paradigm there is little else to look at, which may tempt the viewer to look away from the cued target. In the traditional paradigm, it may often be the actual task of the observers to look at the target, which was not the case for the experiments here or in Hermens and Walker (2015).

5.3 Limitations and future directions

The studies in this thesis have demonstrated that studying social attention in a more ecologically valid method is possible and necessary. The results have shown that cues' effects depend on the exact paradigm and stimuli used. In the real-world, cues are less often looked at, and searching for a target also makes it more likely for the cues to be ignored. Due to the extensive work involved in setting up the studies, recruiting and testing participants and analysing data, not all possibly interesting research questions could be examined. This section will therefore provide some suggestions how this work could be taken forward.

5.3.1 Interfering objects in the scene

One of the main observations in Chapters 2 and 3 studies (employing static and dynamic cues, respectively) is that fixations on the cue only sporadically led to fixations on the target. In fact, all of the studies in these two chapters have shown that most saccades from the cues were directed other elements in the scenes. It is unclear why this was the case, and it would be interesting to further explore possible reasons. Potentially, other items were on the way of the saccades towards the targets, which were distracting the participants into directly looking at them. Future studies could therefore vary the scenes in how many other objects they contain, and how many objects are positioned between the cue and the target. Interesting would also be to examine whether restricting participants' head movements (which improved measurements of gaze direction) influenced gaze behaviour. Possibly this could be done in a desk-based eye tracker without a chin rest, or using a mobile eye tracker, although the latter has the disadvantage that analysing the results is much more labour intensive.

5.3.2 Cues' and targets' Regions of Interest

A possible issue for the laboratory studies in this thesis was the difference in size between cues' and targets' regions of interest. Throughout the experiments in Chapter 2 and 3, cues and targets of different size were used. This might have affected the way participants spent time looking at these elements in the scene. For instance participants might have spent more time looking at actors in the scene (larger size ROIs) and less time on the arrows, as the latter regions of interest were smaller. Similarly, people might have spent less time looking at a smaller targeted item in the scene, compared to a larger targeted item presented in another scene.

The choice to use different targeted items was done for two reasons: (1) to minimize any familiarity from the repetitive use of the same targeted item throughout the scenes and (2) to match the items to the context of the room (e.g., a kettle inside a kitchen). Although, using targets of different characteristics (e.g., size, shape) might have been an issue on the analyses (for the reasons described above), by exposing the targets to all the three cues and by using the same targets per scene, the difference between targets' ROIs can be overcome.

The different size of the cues' ROIs, can also be a limitation for the analyses in the laboratory studies. Social cues were produced by different actors and defined by different body parts (e.g., full body, head and hand for pointing). Arrows used in all the studies were consistent, but still smaller than the two social cues. These differences between the social cues and the arrows as well as the difference between the two social cues (e.g., differences between the individual body parts) could have affected the results (e.g., spend more time looking at larger ROIs). A typical way to address this issue it to normalize the size of cues' ROIs and explore if similar results will occur (e.g., if arrow still fails to attract participants' attention). In this way it can be determined if larger regions of interest received more fixations compared to smaller ones. However, previous studies (e.g., Birmingham et al., 2009; Hermens & Walker, 2015) which compared normalised and not normalised data of social and symbolic cues, have shown similar results (e.g., participants preferring social cues over the arrows). Therefore it seems that size might not be an issue and that participants' attention towards the cues might be influenced by the meaning and the biological relevance cues have.

Although, normalizing the differences between cue' ROIs by size can be a potential solution, there are a lot of issues with this approach. As the two social cues were defined by different body parts it is important to decide which sub-regions of gazing and pointing cues to use, normalize and compare with arrows (e.g., compare just the arm with arrow or face with arrow). Moreover, apart from ROIs size, the location of the cues' ROIs in the scenes, could have an effect in participants attention. It is well known that elements closer to the center of the scene tend to capture participants' attention (e.g., Tatler, 2007). Therefore, people might spend more time looking at a smaller ROI when the latter is positioned closer to the center of the scene compared to a larger ROI positioned away from the center of the scene.

Another way to overcome differences between social and symbolic cues' ROIs size is by increasing the size of the arrow closer to the size of the two social cues (i.e., normalising cues' size inside the images). However, by increasing the size of the arrows might have seemed unusual to the participants and out of context (e.g., a huge arrow seen inside a room). The way arrows used in the studies (i.e., in a form of a sign) in this thesis, is similar to the way arrows are encountered in the real world (i.e., a fairly small size arrow, compared to the size of a person).

Finally, it is important to note that independent of which approach is taken to deal with the differences in cues' ROIs, this analysis is possible (at the moment) only for the static scenes as moving regions of interest are complicated and difficult to analyse.

5.3.3 Controlling scene complexity

Varying the complexity of the scene may be difficult, so another approach would be to ask independent raters to rate the perceived complexity of the scene or to indicate which objects may be of interest to look at. Two recent studies by Henderson and Hayes (2017, 2018) suggested that meaning rather than saliency is what determines people's attention, and it would be interesting to determine whether such results extent to situations where stimuli contain cues of direction. The present study tried to not surround the target with more salient objects, but this only succeeds as far as the experimenter can decide what is salient. For natural scenes it is never entirely possible to control for all confounding variables, but it needs to be realised that in the real-world, such variables will not be controlled for either.

5.3.4 Limited presentation duration

In the second chapter in this thesis it was discussed that the time images were presented might not be sufficient for the gaze cues to develop the cueing effect of the cues. Although, previous studies (e.g., Hermens et al., 2015; Theeuwes & der Stigchel, 2006) have suggested that faces can be detected in just 100ms and that cues and targets are often looked at within the first 1500ms, still images presentation duration might have been a factor influencing participants' attention. Maybe participants looked at the cues just before the offset in the image, not allowing further time to look at the target. Future analyses could therefore examine how long the time between looking at the cue and looking at the target was, and in what proportion of cases the cue was looked at less than this time before the offset of the image. An alternative would be to conduct a further experiment in which the presentation duration of the images is varied. The short presentation durations may also not have given participants enough time to start looking at the cue, but that also means that the cues did not strongly attract people's attention.

5.3.5 Dynamic arrow cues

A further limitation is that the arrows in the dynamic cues experiment were not as dynamic as the social cues in the same experiment. In an attempt to make the arrows dynamic, they were added into the scene at about the same time as the actors made their head turn or pointing gesture. There is, however, quite a big difference between an arrow appearing and an actor, who is already present, start a movement (for further discussion see Chapter 3 Discussion). The reason for letting the arrow appear was to make it visually similar to the arrows used in Chapter 2. Future studies, however, should try alternative arrows, for example, like the one used by (Ristic et al., 2002). They presented a horizontal line, which was later complemented by the arrow's head and tail.

5.3.6 Controlling for motion cues

Only few studies have tried to use dynamic social cues. An important reason is the confound between the effect of motion itself and the cueing nature of the stimuli (e.g.,Farroni et al., 2000). In fact, it has been suggested that motion and direction may be processed by the same neurons (Perrett et al., 1992). Past studies have attempted to compensate for motion effects, by shifting the head in the opposite direction of the eyes inside the head (Bayliss et al., 2005), therefore creating a stronger motion cue away from the gaze shift (produced by the signals from the head) than in the gazed-at direction (the pupils of the eyes). Cueing still followed the eyes in such cases, suggesting gaze cueing overrides motion cues. It was unclear how to achieve such an opposite motion (e.g., head turned right and eyes looking left) for the social and symbolic cues in Chapter 3.

5.3.7 Measuring attention in peripheral vision

The real-world studies of Chapter 4 made use of an eye tracker built into glasses (Tobii 2 glasses) or a head-set (Positive Science eye tracker). Both eye trackers have difficulties recording gaze direction when participants look away from the center of the head direction (i.e., away from the center of the recorded video image). The Positive Science system provides an image of the eye of the participant, and therefore makes coding gaze direction for gaze positions outside the video image possible. Besides extreme eye positions, eye tracker also record missing or incorrect values for other reasons, for example, due to a sudden change in the light intensity in the environment. As a possible solution to these technical limitations, participants could be asked to self-report their gaze direction after the experiment by playing back the recordings to them. Here, however, it needs to be taken into account that participants are not always accurate in recalling their own eye movements (Clarke, Mahon, Irvine, & Hunt, 2017).

5.3.8 Controlling participant expectations

The real-world experiments showed that participants did not make use of the social cues. A possible reason is that participants were too strongly aware that they were part of the experiment and therefore did not trust these cues. Future studies should therefore aim to reduce the effects of participants feeling there are being watched. A common strategy to reduce the sense of being watched is to wait longer between fitting the eye tracker and starting the recording of interest. Another strategy is to observe people's behaviour in an actual setting, but this would not allow for eye movement recordings. It may also help to make the task more realistic. Instead of looking for objects in a room, future studies may ask participants to locate a room. Similar to the methodology used in Chapter 4 (Experiment 6) cues can be provided by actual people and sign which already exist in the navigating space. Here it is worth noticing that the way social cues are used should reflect the way we encounter them in real life search task. For instance, in real life search pointing and gazing can coexist (i.e., same person providing both cues). However, this can be overcome by analysing the order cues are prioritised (e.g., do people first look at the direction of the head and then the hand). In addition, it is important to introduce verbal instructions in the real-world studies to enhance the sense of being part of an actual search. Although it is known that language can overcome cues effects (e.g.,Rayner, Rotello, Stewart, Keir, & Duffy, 2001) however, verbal instructions are an essential element that both social and symbolic cues are encountered with. Again through the analysis, it can be determined at which point, cues are looked at and followed (e.g., after the participants receive verbal instructions about a certain direction). Alongside with how cues will be provided, future real-world studies can manipulate the duration the experiment is conducted. In laboratory settings social attention experiments rarely last for a long time and cues are presented for a brief time. What we can suggest is that ideally real-world experiments should be more flexibly in terms of time to see if a more natural gaze behaviour and attention will emerge.

5.4 Conclusion

This thesis has studied where observers look when confronted with social (people pointing and gazing) and symbolic cues (arrows) in their environment. Three main scenarios were studied: (1) viewing static images (photographs), (2) viewing dynamic scenes (videos), (3) real-world viewing. When static images were used, observers fixated the social cues more often than the symbolic cues. Cues were not always fixated and when cues were fixated, this did not often lead to subsequent gazes on the looked-at or pointed-at object. For dynamic cues, the social cues attracted more attention, but did not often lead to gaze following. In the real-world situation, where participants were asked to find a target object, cues were not often looked at and did not substantially aid in finding the target object. These results contrast the strong effects of gaze and pointing cues found in more traditional social attention experiments, where cues are presented at fixation and in isolation. The results from this thesis therefore suggest that social attention studied in the more traditional paradigm may not substantially aid the understanding of social attention in the real-world. The experiments also suggest that studying social attention in the real-world is possible, although more time-consuming and complex experiments are needed.

References

Akiyama, T., Kato, M., Muramatsu, T., Saito, F., Umeda, S., & Kashima, H. (2006). Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. *Neuropsychologia*, *44*(10), 1804–1810. doi: 10.1016/j.neuropsychologia.2006.03.007

Allison, P. D. (2010). *Survival analysis using sas: a practical guide*. Sas Institute.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278. doi: 10.1016/s1364-6613(00)01501-1

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 997–1010. doi: 10.1037/0278-7393.25.4.997

Anderson, N. C., Risko, E. F., & Kingstone, A. (2011). Exploiting human sensitivity to gaze for tracking the eyes. *Behavior Research Methods*, *43*(3), 843–852. doi: 10.3758/s13428-011-0078-8

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *The American Journal of Psychology*, *82*(4), 474. doi: 10.2307/1420441

Argyle, M., & Cook, M. (1976). Gaze and mutual gaze. *Cambridge U Press*.

Azarian, B., Buzzell, G. A., Esser, E. G., Dornstauder, A., & Peterson, M. S. (2017). Averted body postures facilitate orienting of the eyes. *Acta Psychologica*, *175*, 28–32. doi: 10.1016/j.actpsy.2017.02.006

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390–412.

Barnes, K. A., Kaplan, L. A., & Vaidya, C. J. (2007). Developmental differences in cognitive control of socio-affective processing. *Developmental Neuropsychology*, *32*(3), 787–807. doi: 10.1080/87565640701539576

Baron-Cohen, S. (1995). The eye direction detector (edd) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore & e. Dunham P.J. (Eds.), *Joint attention: Its origins and role in development* (p. 71-79). Lawrence Erlbaum Associates, Inc.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*(3), 412–414. doi: 10.1098/rsbl.2006.0509

Bayliss, A. P., di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *The Quarterly Journal of Experimental Psychology Section A*, *58*(4), 631–650. doi: 10.1080/02724980443000124

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644–653.

Bayliss, A. P., & Tipper, S. P. (2006a). Gaze cues evoke both spatial and object-centered shifts of attention. *Perception & Psychophysics*, *68*(2), 310–318.

Bayliss, A. P., & Tipper, S. P. (2006b). Predictive gaze cues and personality judgments. *Psychological Science*, *17*(6), 514–520. doi: 10.1111/j.1467-9280.2006.01737.x

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, Y. D. (2002). Priming visual face-processing mechanisms: Electrophysiological evidence. *Psychological Science*, *13*(2), 190–193. doi: 10.1111/1467-9280.00435

Bindemann, M., Burton, A. M., Hooge, I. T. C., Jenkins, R., & de Haan, E. H. F. (2005). Faces retain attention. *Psychonomic Bulletin & Review*, *12*(6), 1048–1053. doi: 10.3758/bf03206442

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, *9*(2), 7–7. doi: 10.1167/9.2.7

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008a). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology*, *61*(7), 986–998. doi: 10.1080/17470210701410375

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008b). Why do we look at people's eyes? *Journal of Eye Movement Research*, *1*(1), 1–6. doi: https://doi.org/10.16910/jemr.1.1.1

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Get real! resolving the debate about equivalent social stimuli. *Visual Cognition*, *17*(6-7), 904–924. doi: 10.1080/13506280902758044

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, *16*(11), 3737–3744. doi: 10.1523/jneurosci.16-11-03737.1996

Borji, A., & Itti, L. (2014). Defending yarbus: Eye movements reveal observers' task. *Journal of Vision*, *14*(3), 29–29. doi: 10.1167/14.3.29

Brignani, D., Guzzon, D., Marzi, C., & Miniussi, C. (2009). Attentional orienting induced by arrows and eye-gaze compared with an endogenous cue. *Neuropsychologia*, *47*(2), 370–381. doi: 10.1016/j.neuropsychologia.2008.09.011

Burton, A. M., Bindemann, M., Langton, S. R., Schweinberger, S. R., & Jenkins, R. (2009). Gaze perception requires focused attention: Evidence from an interference task. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 108–118. doi: 10.1037/0096-1523.35.1.108

Buswell, G. T. (1935). How people look at pictures: a study of the psychology and perception in art. *Univ. Chicago Press*.

Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, *3*(3), 253–272. doi: 10.1177/016502548000300303

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. doi: 10.1111/j.2044-835x.1991.tb00862.x

Campbell, R., Heywood, C., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, *28*(11), 1123–1142. doi: 10.1016/0028-3932(90)90050-x

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i. doi: 10.2307/1166214

Clarke, A. D. F., Mahon, A., Irvine, A., & Hunt, A. R. (2017). People are unable to recognize or report on their own eye movements. *Quarterly Journal of Experimental Psychology*, *70*(11), 2251–2270. doi: 10.1080/17470218.2016.1231208

Cline, M. G. (1967). The perception of where a person is looking. *The American Journal of Psychology*, *80*(1), 41. doi: 10.2307/1420539

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences*, *95*(3), 831–838. doi: 10.1073/pnas.95.3.831

Crostella, F., Carducci, F., & Aglioti, S. M. (2009). Reflexive social attention is mapped according to effector-specific reference systems. *Experimental Brain Research*, *197*(2), 143–151. doi: 10.1007/s00221-009-1900-8

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141–168. doi: 10.1111/j.1468-0017.2009.01384.x

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, *36*(4), 511–523. doi: 10.1037/0012-1649.36.4.511

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6-7), 790–811. doi: 10.1080/13506280902793843

D'Entremont, B., Hains, S., & Muir, D. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. doi: 10.1016/s0163-6383(97)90048-5

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*(6651), 596–599. doi: 10.1038/39309

Doumen, M. J. A., Kappers, A. M. L., & Koenderink, J. J. (2010). Effects of context on a 3d pointing task. *Journal of Vision*, *6*(6), 728–728. doi: 10.1167/6.6.728

Downing, P., Dodds, C., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition*, *11*(1), 71–79. doi: 10.1080/13506280344000220

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509–540. doi: 10.1080/135062899394920

Dukewich, K. R., Klein, R. M., & Christie, J. (2008). The effect of gaze on gaze direction while looking at art. *Psychonomic Bulletin & Review*, *15*(6), 1141–1147. doi: 10.3758/pbr.15.6.1141

Eacott, M., Heywood, C., Gross, C., & Cowey, A. (1993). Visual discrimination impairments following lesions of the superior temporal sulcus are not specific for facial stimuli. *Neuropsychologia*, *31*(6), 609–619. doi: 10.1016/0028-3932(93)90055-5

Efron, D. (1941). *Gesture and environment: A tentative study of some of the spatio-temporal and" linguistic" aspects of the gestural behavior of eastern jews and southern italians in new york city, living under similar as well as different environmental conditions*. King's crown Press.

Efron, D. (1972). *Gesture, race and culture (The Hague: Mouton)*. Orig-i.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, *22*(4), 353–374. doi: 10.1111/j.1460-2466.1972.tb00163.x

Ellsworth, P. C. (1975). Direct gaze as a social stimulus: The example of aggression. In *Nonverbal communication of aggression* (pp. 53–75). Springer.

Emery, N. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, *24*(6), 581–604. doi: 10.1016/s0149-7634(00)00025-7

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718. doi: 10.1080/13506280050144399

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39–60. doi: 10.1207/s15327078in0501_2

Feyereisen, P., & Lannoy, J.-D. D. (1991). *Gestures and speech: Psychological investigations*. Cambridge University Press.

Firestone, A., Turk-Browne, N. B., & Ryan, J. D. (2007). Age-related deficits in face recognition are related to underlying changes in scanning behavior. *Aging, Neuropsychology, and Cognition*, *14*(6), 594–607. doi: 10.1080/13825580600899717

Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*(4), 571–583. doi: 10.1068/p5705

Fletcher-Watson, S., Leekam, S., Benson, V., Frank, M., & Findlay, J. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, *47*(1), 248–257. doi: 10.1016/j.neuropsychologia.2008.07.016

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319–331.

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? social presence, eye contact and autistic traits. *PloS one*, *8*(1), e53286.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495. doi: 10.3758/bf03208827

Friesen, C. K., & Kingstone, A. (2003). Abrupt onsets and gaze direction cues trigger independent reflexive attentional effects. *Cognition*, *87*(1), B1–B10. doi: 10.1016/s0010-0277(02)00181-6

Friesen, C. K., Moore, C., & Kingstone, A. (2005). Does gaze direction really trigger a reflexive shift of spatial attention? *Brain and Cognition*, *57*(1), 66–69. doi: 10.1016/j.bandc.2004.08.025

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 319.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. doi: 10.1037/0033-2909.133.4.694

Gallup, A. C., Chong, A., & Couzin, I. D. (2012). The directional flow of visual information transfer between pedestrians. *Biology Letters*, *8*(4), 520–522. doi: 10.1098/rsbl.2012.0160

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences*, *109*(19), 7245–7250. doi: 10.1073/pnas.1116141109

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage*, *13*(6), 1102–1112. doi: 10.1006/nimg.2001.0769

Gibson, B. S., & Bryant, T. A. (2005). Variation in cue duration reveals top-down modulation of involuntary orienting to uninformative symbolic cues. *Perception & psychophysics*, *67*(5), 749–758. doi: 10.3758/bf03193530

Gibson, B. S., & Kingstone, A. (2006). Visual attention and the semantics of space: Beyond central and peripheral cues. *Psychological Science*, *17*(7), 622–627. doi: 10.1111/j.1467-9280.2006.01754.x

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *The American journal of psychology*, *76*(3), 386–394.

Greenhouse, J. B., Stangl, D., & Bromberg, J. (1989). An introduction to survival analysis: Statistical methods for analysis of clinical trial data. *Journal of Consulting and Clinical Psychology*, *57*(4), 536.

Gregory, N. J., & Antolin, J. V. (2019). Does social presence or the potential for interaction reduce social gaze in online social scenarios? introducing the "live lab" paradigm. *Quarterly Journal of Experimental Psychology*, *72*(4), 779–791.

Gregory, N. J., Hermens, F., Facey, R., & Hodgson, T. L. (2016). The developmental trajectory of attentional orienting to socio-biological cues. *Experimental Brain Research*, *234*(6), 1351–1362. doi: 10.1007/s00221-016-4627-3

Grèzes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain*, *122*(10), 1875–1887. doi: 10.1093/brain/122.10.1875

Guo, K. (2012). Holistic gaze strategy to categorize facial expression of varying intensities. *PLoS ONE*, *7*(8), e42585. doi: 10.1371/journal.pone.0042585

Guzzon, D., Brignani, D., Miniussi, C., & Marzi, C. (2010). Orienting of attention with eye and arrow cues and the effect of overtraining. *Acta Psychologica*, *134*(3), 353–362. doi: 10.1016/j.actpsy.2010.03.008

Hayward, D., & Ristic, J. (2018). Changes in tonic alertness but not voluntary temporal preparation modulate the attention elicited by task-relevant gaze and arrow cues. *Vision*, *2*(2), 18. doi: 10.3390/vision2020018

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, *1*(10), 743–747. doi: 10.1038/s41562-017-0208-0

Henderson, J. M., & Hayes, T. R. (2018). Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *Journal of Vision*, *18*(6), 10. doi: 10.1167/18.6.10

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*(1), 98–106. doi: 10.3758/bf03195300

Hermens, F. (2015). Fixation instruction influences gaze cueing. *Visual Cognition*, *23*(4), 432–449. doi: 10.1080/13506285.2015.1042539

Hermens, F. (2017). The effects of social and symbolic cues on visual search: Cue shape trumps biological relevance. *Psihologija*, *50*(2), 117–140. doi: 10.2298/psi161005003h

Hermens, F., Bindemann, M., & Burton, A. M. (2015). Responding to social and symbolic extrafoveal cues: cue shape trumps biological relevance. *Psychological Research*, *81*(1), 24–42. doi: 10.1007/s00426-015-0733-2

Hermens, F., & Gielen, S. (2003). Catching oriented objects. *Acta Psychologica*, *114*(1), 17–39. doi: 10.1016/s0001-6918(03)00048-9

Hermens, F., Golubickis, M., & Macrae, C. N. (2018). Eye movements while judging faces

for trustworthiness and dominance. *PeerJ*, *6*, e5702.
Hermens, F., & Walker, R. (2010). Gaze and arrow distractors influence saccade trajectories similarly. *Quarterly Journal of Experimental Psychology*, *63*(11), 2120–2140. doi: 10.1080/17470211003718721
Hermens, F., & Walker, R. (2012). Do you look where i look? Attention shifts and response preparation following dynamic social cues. *Journal of Eye Movement Research,*, *5*(5), 1–11. doi: https://doi.org/10.16910/jemr.5.5.5
Hermens, F., & Walker, R. (2015). The influence of social and symbolic cues on observers' gaze behaviour. *British Journal of Psychology*, *107*(3), 484–502. doi: 10.1111/bjop.12159
Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *Neuroreport*, *10*(16), 3443–3447.
Hietanen, J. K. (2002). Social attention orienting integrates visual information from head and body orientation. *Psychological Research*, *66*(3), 174–179.
Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage*, *33*(1), 406–413. doi: 10.1016/j.neuroimage.2006.06.048
Hill, J. L., Patel, S., Gu, X., Seyedali, N. S., Bachevalier, J., & Sereno, A. B. (2010). Social orienting: reflexive versus voluntary control. *Vision research*, *50*(20), 2080–2092.
Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–84. doi: 10.1038/71152
Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 58–69. doi: 10.1037/0278-7393.32.1.58
Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological science*, *12*(5), 360–365. doi: 10.1111/1467-9280.00367
Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134. doi: 10.1111/1467-9280.00024
Hooker, C. I., Paller, K. A., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., & Reber, P. J. (2003). Brain networks for analyzing eye gaze. *Cognitive Brain Research*, *17*(2), 406–418. doi: 10.1016/s0926-6410(03)00143-5
Ioannidou, F., Hermens, F., & Hodgson, T. L. (2016). The central bias in day-to-day viewing. *Journal of Eye Movement Research*, *9*(6), 1–13. doi: 10.16910/JEMR.9.6.6
Itier, R. J., Villate, C., & Ryan, J. D. (2007). Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia*, *45*(5), 1019–1028. doi: 10.1016/j.neuropsychologia.2006.09.004
Ivanoff, J., & Saoud, W. (2009). Nonattentional effects of nonpredictive central cues. *Attention, Perception, & Psychophysics*, *71*(4), 872–880. doi: 10.3758/app.71.4.872
Jakobsen, K. V., Frick, J. E., & Simpson, E. A. (2013). Look here! the development of attentional orienting to symbolic cues. *Journal of Cognition and Development*, *14*(2), 229–249. doi: 10.1080/15248372.2012.666772
Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In I. J.Long & A. D. Baddeley (Eds.), *Attention and performance* (pp. 187,203). Hillsdale, NJ:Erlbaum.
Keiley, M. K., & Martin, N. C. (2005). Survival analysis in family research. *Journal of Family Psychology*, *19*(1), 142–156. doi: 10.1037/0893-3200.19.1.142

Kendon, A. (1986). Current issues in the study of gesture. *The biological foundations of gestures: Motor and semiotic aspects*, *1*, 23–47.

Kendon, A. (1994). Introduction to the special issue: Gesture and understanding in interaction. *Research on Language & Social Interaction*, *27*(3), 171–173. doi: 10.1207/s15327973rlsi2703_1

Kingstone, A., Friesen, C. K., & Gazzaniga, M. S. (2000). Reflexive joint attention depends on lateralized cortical connections. *Psychological Science*, *11*(2), 159–166. doi: 10.1111/1467-9280.00232

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*(5), 176–180. doi: 10.1111/1467-8721.01255

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it!: An fMRI investigation. *Brain and Cognition*, *55*(2), 269–271. doi: 10.1016/j.bandc.2004.02.037

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological bulletin*, *100*(1), 78.

Kuhn, G., & Benson, V. (2007). The influence of eye-gaze and arrow pointing distractor cues on voluntary eye movements. *Perception & Psychophysics*, *69*(6), 966–971. doi: 10.3758/bf03193934

Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception & Psychophysics*, *71*(2), 314–327. doi: 10.3758/app.71.2.314

Kuhn, G., Pagano, A., Maani, S., & Bunce, D. (2015). Age-related decline in the reflexive component of overt gaze following. *Quarterly Journal of Experimental Psychology*, *68*(6), 1073–1081. doi: 10.1080/17470218.2014.975257

Kuhn, G., Tatler, B. W., & Cole, G. G. (2009). You look where i look! Effect of gaze cues on overt and covert attention in misdirection. *Visual Cognition*, *17*(6-7), 925–944. doi: 10.1080/13506280902826775

Kuhn, G., & Tipples, J. (2010). Increased gaze following for fearful faces. it depends on what you're looking for! *Psychonomic Bulletin & Review*, *18*(1), 89–95. doi: 10.3758/s13423-010-0033-1

Lai, H., Li, F., & Wechsler, H. (2007). Robust face recognition strategies using feed-forward architectures and parts. In *International workshop on analysis and modeling of faces and gestures* (pp. 290–304).

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences*, *108*(14), 5548–5553. doi: 10.1073/pnas.1017022108

Land, M., & Tatler, B. (2009). *Looking and acting: vision and eye movements in natural behaviour*. Oxford University Press.

Langton., S. R. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(3), 825–845.

Langton, S. R., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*(5), 541–567. doi: 10.1080/135062899394939

Langton, S. R., & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 747.

Langton, S. R., O'Malley, C., & Bruce, V. (1996). Actions speak no louder than words:

Symmetrical cross-modal interference effects in the processing of verbal and gestural information. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1357–1375. doi: 10.1037/0096-1523.22.6.1357

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59. doi: 10.1016/s1364-6613(99)01436-9

Lempers, J. D. (1976). *Production of pointing, comprehension of pointing and understanding of looking behavior in young children.* (Unpublished doctoral dissertation). ProQuest Information & Learning.

Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, *17*(2), 215–220. doi: 10.1037/0012-1649.17.2.215

Levy, J., Foulsham, T., & Kingstone, A. (2012). Monsters are people too. *Biology Letters*, *9*(1), 20120850–20120850.

Liszkowski, U. (2005). Human twelve-month-olds point cooperatively to share interest with and helpfully provide information for a communicative partner. *Gestural Communication in Nonhuman and Human Primates*, *5*(1-2), 135–154. doi: 10.1075/gest.5.1.11lis

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*(3), 297–307. doi: 10.1111/j.1467-7687.2004.00349.x

Ludwig, C. J. H., Ranson, A., & Gilchrist, I. D. (2008). Oculomotor capture by transient events: A comparison of abrupt onsets, offsets, motion, and flicker. *Journal of Vision*, *8*(14), 11–11. doi: 10.1167/8.14.11

MacDonald, R., & Tatler, B. (2013). The effect of social roles on gaze cue utilisation in a real-world collaboration. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*(35).

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of vision*, *11*(9), 9–9.

MacLeod, C. M. (1991). Half a century of research on the stroop effect: an integrative review. *Psychological bulletin*, *109*(2), 163.

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (lmmgui). *Frontiers in psychology*, *6*, 2.

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11), 8–8. doi: 10.1167/9.11.8

Maurer, D. (1985). Infants' perception of facedness. *Social perception in infants*. Retrieved from `https://ci.nii.ac.jp/naid/10021891245/en/`

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350–371. doi: 10.1037/0033-295x.92.3.350

McNeill, D. (1987). So you do think gestures are nonverbalp? Reply to Feyereisen (1987). *Psychological Review*, *94*(4), 499–504. doi: 10.1037/0033-295x.94.4.499

McNeill, D. (1989). A straight path—to where? Reply to Butterworth and Hadar. *Psychological Review*, *96*(1), 175–179.

Miller Jr, R. G. (2011). *Survival analysis* (Vol. 66). John Wiley & Sons.

Morissette, P., Ricard, M., & Décarie, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology*, *13*(2), 163–175. doi: 10.1111/j.2044-835x.1995.tb00671.x

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 615–622. doi: 10.1037/0278-7393.18.3.615

Nakamura, A., Maess, B., Knösche, T. R., Gunter, T. C., Bach, P., & Friederici, A. D. (2004). Cooperation of different neuronal systems during hand sign recognition. *NeuroImage*, *23*(1), 25–34. doi: 10.1016/j.neuroimage.2004.04.034

Nummenmaa, L., & Hietanen, J. K. (2009). How attentional systems process conflicting cues. the superiority of social over symbolic orienting revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1738–1754. doi: 10.1037/a0016472

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, *6*(2), 99–116. doi: 10.1162/jocn.1994.6.2.99

Orquin, J. L., Ashby, N. J. S., & Clarke, A. D. F. (2015). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*, *29*(2-3), 103–115. doi: 10.1002/bdm.1867

Panis, S., & Hermens, F. (2014). Time course of spatial contextual interference: Event history analyses of simultaneous masking by nonoverlapping patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 129–144. doi: 10.1037/a0032949

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, *16*(10), 1706–1716. doi: 10.1162/0898929042947900

Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, *41*(2), 156–170. doi: 10.1016/s0028-3932(02)00146-x

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical transactions of the royal society of London. Series B: Biological sciences*, *335*(1273), 23–30.

Perrett, D. I., Oram, M., Hietanen, J., & Benson, P. (1994). Issues of representation in object vision. *The neuropsychology of high-level vision*, 33–61.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *223*(1232), 293–317. doi: 10.1098/rspb.1985.0003

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi: 10.1080/00335558008248231

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. *Modes of perceiving and processing information*, *137*(158), 2.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience*, *18*(6), 2188–2199. doi: 10.1523/jneurosci.18-06-02188.1998

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*(4), 413–425.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading*.

Psychology Press.
Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, *7*(3), 219–226. doi: 10.1037/1076-898x.7.3.219
Reuter-Lorenz, P., Oonk, H., Barnes, L., & Hughes, H. (1995). Effects of warning signals and fixation point offsets on the latencies of pro- versus antisaccades: implications for an interpretation of the gap effect. *Experimental Brain Research*, *103*(2). doi: 10.1007/bf00231715
Ricciardelli, P., Betta, E., Pruner, S., & Turatto, M. (2009). Is there a direct link between gaze perception and joint attention behaviours? Effects of gaze contrast polarity on oculomotor behaviour. *Experimental Brain Research*, *194*(3), 347–357. doi: 10.1007/s00221-009-1706-8
Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *Neuroreport*, *13*(17), 2259–2264.
Rimé, B., & Schiaratura, L. (1991). Gesture and speech. *Fundamentals of nonverbal behavior*, 239–281.
Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*(3), 507–513. doi: 10.3758/bf03196306
Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention. *Cognition*, *94*(3), 55–65. doi: 10.1016/j.cognition.2004.04.005
Ristic, J., & Kingstone, A. (2006). Attention to arrows: Pointing to a new direction. *The Quarterly Journal of Experimental Psychology*, *59*(11), 1921–1930. doi: https://doi.org/10.1080/17470210500416367
Ristic, J., Wright, A., & Kingstone, A. (2007). Attentional control and reflexive orienting to gaze and arrow cues. *Psychonomic Bulletin & Review*, *14*(5), 964–969. doi: 10.3758/bf03194129
Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*(3), 426–435. doi: 10.1111/j.1467-7687.2012.01139.x
Rutherford, M. D., & Krysko, K. M. (2008). Eye direction, not movement direction, predicts attention shifts in those with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(10), 1958–1965. doi: 10.1007/s10803-008-0592-4
Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2008). Time course of superior temporal sulcus activity in response to eye gaze: a combined fMRI and MEG study. *Social Cognitive and Affective Neuroscience*, *3*(3), 224–232. doi: 10.1093/scan/nsn016
Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2009). Commonalities in the neural mechanisms underlying automatic attentional shifts by gaze, gestures, and symbols. *NeuroImage*, *45*(3), 984–992. doi: 10.1016/j.neuroimage.2008.12.052
Sato, W., Okada, T., & Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. *Experimental Brain Research*, *183*(1), 87–94. doi: 10.1007/s00221-007-1025-x
Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*(5), 402–409. doi: 10.1111/1467-9280.00472
Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, *12*(1), 127–144. doi: 10.1080/13506280444000157
Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis

when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin*, *110*(2), 268–290. doi: 10.1037/0033-2909.110.2.268

Singer, J. D., & Willett, J. B. (2003). Survival analysis. *Handbook of psychology*, 555–580.

Smilek, D., Birmingham, E., Cameron, D., Bischof, W., & Kingstone, A. (2006). Cognitive ethology and exploring attention in real-world scenes. *Brain Research*, *1080*(1), 101–119. doi: 10.1016/j.brainres.2005.12.090

Stins, J. F., Polderman, J. C. T., Boomsma, D. I., & de Geus, E. J. C. (2007). Conditional accuracy in response interference tasks: Evidence from the eriksen flanker task and the spatial conflict task. *Advances in Cognitive Psychology*, *3*(3), 409–417. doi: 10.2478/v10053-008-0005-4

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*(6), 495–508. doi: 10.3758/bf03212183

Sun, H.-M., & Balas, B. (2014). Face features and face configurations both contribute to visual crowding. *Attention, Perception, & Psychophysics*, *77*(2), 508–519. doi: 10.3758/s13414-014-0786-0

Swettenham, J., Condie, S., Campbell, R., Milne, E., & Coleman, M. (2003). Does the perception of moving eyes trigger reflexive visual orienting in autism? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 325–334. doi: 10.1098/rstb.2002.1203

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 4. doi: 10.1167/7.14.4

Tatler, B. W., & Tatler, S. L. (2013). The influence of instructions on object memory in a real-world setting. *Journal of vision*, *13*(2), 5–5.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, *1*(1), 7–27. doi: 10.1068/i0382

Theeuwes, J., & der Stigchel, S. V. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, *13*(6), 657–665. doi: 10.1080/13506280500410949

Therneau, T. M. (2015). A package for survival analysis in s [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=survival` (version 2.38)

Thoermer, C., & Sodian, B. (2001). Preverbal infants' understanding of referential gestures. *First Language*, *21*(63), 245–264.

Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *NeuroImage*, *37*(3), 966–973. doi: 10.1016/j.neuroimage.2007.05.058

Thompson, L. A., & Massaro, D. W. (1986). Evaluation and integration of speech and pointing gestures during referential understanding. *Journal of Experimental Child Psychology*, *42*(1), 144–168. doi: 10.1016/0022-0965(86)90020-2

Thompson, L. A., & Massaro, D. W. (1994). Children′s integration of speech and pointing gestures in comprehension. *Journal of Experimental Child Psychology*, *57*(3), 327–354. doi: 10.1006/jecp.1994.1016

Tipper, C. M., Handy, T. C., Giesbrecht, B., & Kingstone, A. (2008). Brain responses to biological relevance. *Journal of Cognitive Neuroscience*, *20*(5), 879–891. doi: 10.1162/jocn.2008.20510

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*(2), 314–318. doi: 10.3758/bf03196287

Tipples, J. (2008). Orienting to counterpredictive gaze and arrow cues. *Perception & Psychophysics*, *70*(1), 77–87. doi: 10.3758/pp.70.1.77

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, *52*(3), 314–320. doi: 10.1016/j.jhevol.2006.10.001

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. doi: 10.1037/0033-295x.113.4.766

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. *Cognitive Science*, *28*(2), 289–301. doi: 10.1207/s15516709cog2802_8

Võ, M. L.-H. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision*, *10*(3), 1–13. doi: 10.1167/10.3.14

Võ, M. L. H., & Henderson, J. M. (2009). Does gravity matter? effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, *9*(3), 24–24. doi: 10.1167/9.3.24

Walker, R., Deubel, H., Schneider, W. X., & Findlay, J. M. (1997). Effect of remote distractors on saccade programming: Evidence for an extended fixation zone. *Journal of Neurophysiology*, *78*(2), 1108–1119. doi: 10.1152/jn.1997.78.2.1108

Wicker, B., Michel, F., Henaff, M.-A., & Decety, J. (1998). Brain regions involved in the perception of gaze: A PET study. *NeuroImage*, *8*(2), 221–227. doi: 10.1006/nimg.1998.0357

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77–84. doi: 10.1016/j.tics.2010.12.001

Yarbus, A. L. (1967). *Eye movements and vision*. Springer US. doi: 10.1007/978-1-4899-5379-7

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835. doi: 10.1037/a0013118

Zwickel, J., & Võ, M. L.-H. (2010). How the presence of persons biases eye movements. *Psychonomic Bulletin & Review*, *17*(2), 257–262.

Appendix A

Appendix

A.1 Examples of R codes used in Chapters 2, 3 and 4

A.1.1 Code exploring the main effect of the cues (Experiment 1, Chapter 2)

```
With main effect <-DwellTimes ~ Cues + (1|Subjects) +
                  (1|Images), data= DTCues
Without main effect <- DwellTimes ~ (1|Subjects) +
                  (1|Images), data=DTCues
```

A.1.2 Code exploring a two-way interaction between cues and task (Experiment 4, Chapter 3)

```
Model1 <- DwellTimes ~ Cues * Task + (1|Subjects) +
          (1|Images), data = DTCues


Model 2 <- DwellTimes ~ Cues + Task + (1|Subjects) +
          (1|Images), data = DTCues
```

A.2 Cues' Region of Interest normalized size, Chapter 2.

To normalise ROI's size we counted how many pixels are under each ROI (ROIs area) and then we divided this ROIs area over the overall area (overall number of pixels) in the picture. This gave a normalised value between 0 and 1 which represents the area as a fraction of the overall image size. Please note that the values for Experiment 2 are not the individual values for each cueing condition but the mean value from the combination these cues were encountered (e.g., an arrow cue with a pointing cue).

Table A.1 Cues' Region of Interest normalized size per cueing condition and experiment in Chapter 2.

Experiment	Cue	ROI	Normalized size
Experiment 1	Arrow	Arrow	0.01
	Gaze	Head	0.01
		Body	0.1
	Point	Head	0.01
		Body	0.1
		Arm	0.01
Experiment 2	Arrow - Point	Arrow	0.01
		Head	0.03
		Body	0.2
		Arm	0.02
	Arrow - Gaze	Arrow	0.01
		Head	0.03
		Body	0.1
	Gaze - Point	Head (Gaze)	0.03
		Body (Gaze)	0.2
		Head (Point)	0.03
		Body (Point)	0.2
		Arm (Point)	0.02
Experiment 3	Arrow	Arrow	0.01
	Gaze	Head	0.01
		Body	0.1
	Point	Head	0.01
		Body	0.1
		Arm	0.01

A.3 Results for the percentage of trials with fixations on the cues and targets in Experiment 2, Chapter 2.

Table A.2 Mixed effect statistics comparing different cue combinations for the percentages on trials with fixations on the cues in Experiment 2, Chapter 2.

Combination 1	Combination 2	χ^2-value	p-value
Arrow - Gaze away	Arrow - Point away	2.45	0.12
Arrow - Gaze away	Gaze – Arrow away	32.86	<0.0001
Arrow - Gaze away	Gaze – Point away	5.02	0.03
Arrow - Gaze away	Point – Arrow away	31.36	<0.0001
Arrow - Gaze away	Point – Gaze away	10.71	0.001
Arrow – Point away	Gaze – Arrow away	45.51	<0.0001
Arrow – Point away	Gaze – Point away	14.07	0.0002
Arrow – Point away	Point – Arrow away	44.49	<0.0001
Arrow – Point away	Point – Gaze away	20.59	<0.0001
Gaze – Arrow away	Gaze – Point away	14.61	<0.0001
Gaze – Arrow away	Point – Arrow away	0.14	0.71
Gaze – Arrow away	Point – Gaze away	8.62	0.003
Gaze – Point away	Point – Arrow away	12.53	0.0004
Gaze – Point away	Point – Gaze away	1.11	0.29
Point – Arrow away	Point – Gaze away	6.42	0.01

Table A.3 Mixed effect statistics comparing different cue combinations for the percentages on trials with fixations on the target in Experiment 2, Chapter 2.

Combination 1	Combination 2	χ^2-value	p-value
Arrow - Gaze away	Arrow - Point away	0.58	0.45
Arrow - Gaze away	Gaze – Arrow away	1.01	0.31
Arrow - Gaze away	Gaze – Point away	8.40	0.004
Arrow - Gaze away	Point – Arrow away	6.05	0.01
Arrow - Gaze away	Point – Gaze away	2.86	0.09
Arrow – Point away	Gaze – Arrow away	0.09	0.76
Arrow – Point away	Gaze – Point away	2.46	0.12
Arrow – Point away	Point – Arrow away	9.40	0.002
Arrow – Point away	Point – Gaze away	5.0	0.03
Gaze – Arrow away	Gaze – Point away	3.78	0.05
Gaze – Arrow away	Point – Arrow away	12.94	0.0003
Gaze – Arrow away	Point – Gaze away	7.56	0.006
Gaze – Point away	Point – Arrow away	28.72	<0.0001
Gaze – Point away	Point – Gaze away	18.84	<0.0001
Point – Arrow away	Point – Gaze away	0.90	0.34
No Cue	Arrow - Gaze away	77.65	<0.0001
No Cue	Arrow - Point away	77.88	<0.0001
No Cue	Gaze – Arrow away	92.88	<0.0001
No Cue	Gaze – Point away	134.92	<0.0001
No Cue	Point – Arrow away	53.38	<0.0001
No Cue	Point – Gaze away	59.24	<0.0001

A.4 Dwell times on head, body and arm per actor position in Experiment 3, Chapter 2.

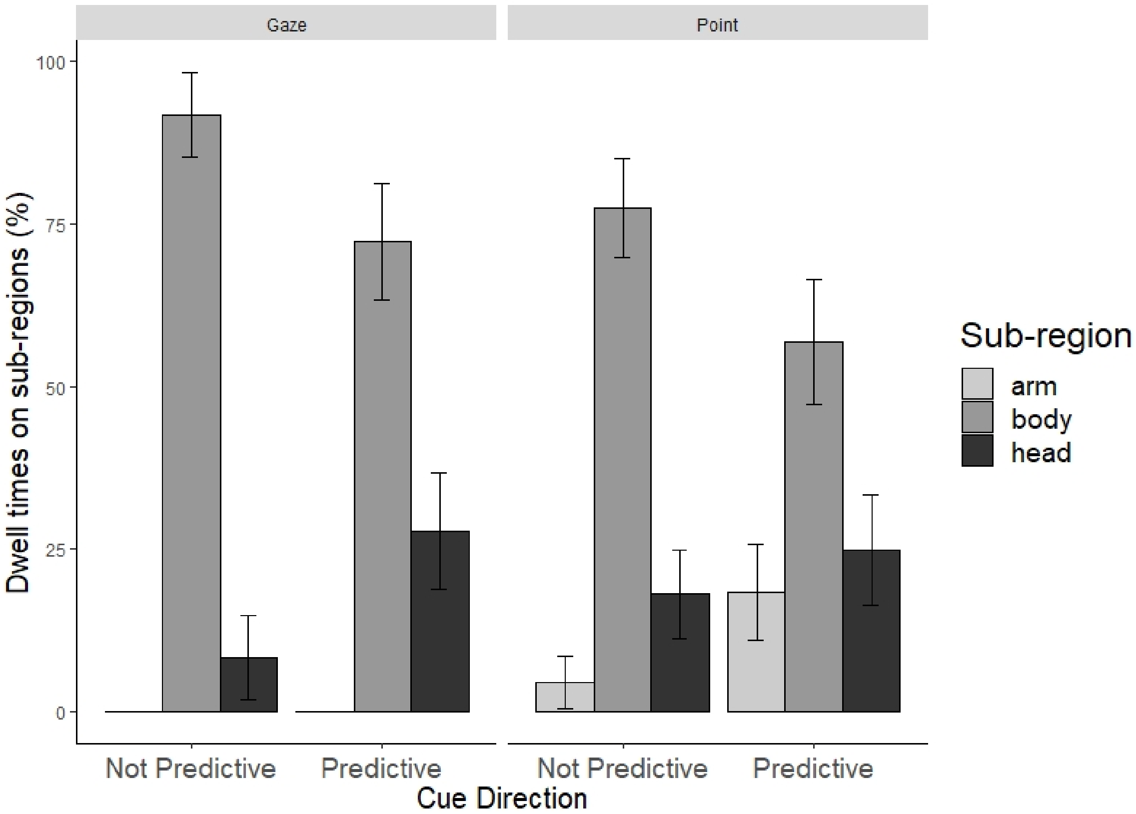

Fig. A.1 Dwell Times on the three body parts (different bars) for the two social cues (gaze cue on the left, pointing cue on the right) and the two cue directions when actors' position was closer to the center of the scene.

Table A.4 Statistics (chi-square value, p-value) for comparisons of dwell times on the three sub-regions per validity and actors' position for the two social cues of Experiment 3, Chapter 2.

Comparison	Cue	Position	Validity	χ^2-value	p-value
Body vs Head	Gaze	Central	Not Predict	50.15	<0.0001
Body vs Head	Gaze	Central	Predict	40.02	<0.0001
Body vs Arm	Point	Central	Not Predict	62.01	<0.0001
Body vs Head				31.85	<0.0001
Body vs Arm	Point	Central	Predict	13.05	0.0003
Body vs Head				5.07	0.02
Body vs Head	Gaze	Periphery	Not Predict	125.45	<0.0001
Body vs Head	Gaze	Periphery	Predict	186.77	<0.0001
Body vs Arm	Point	Periphery	Not Predict	18.26	<0.0001
Body vs Head				68.05	<0.0001
Body vs Arm	Point	Periphery	Predict	42.05	<0.0001
Body vs Head				85.59	<0.0001

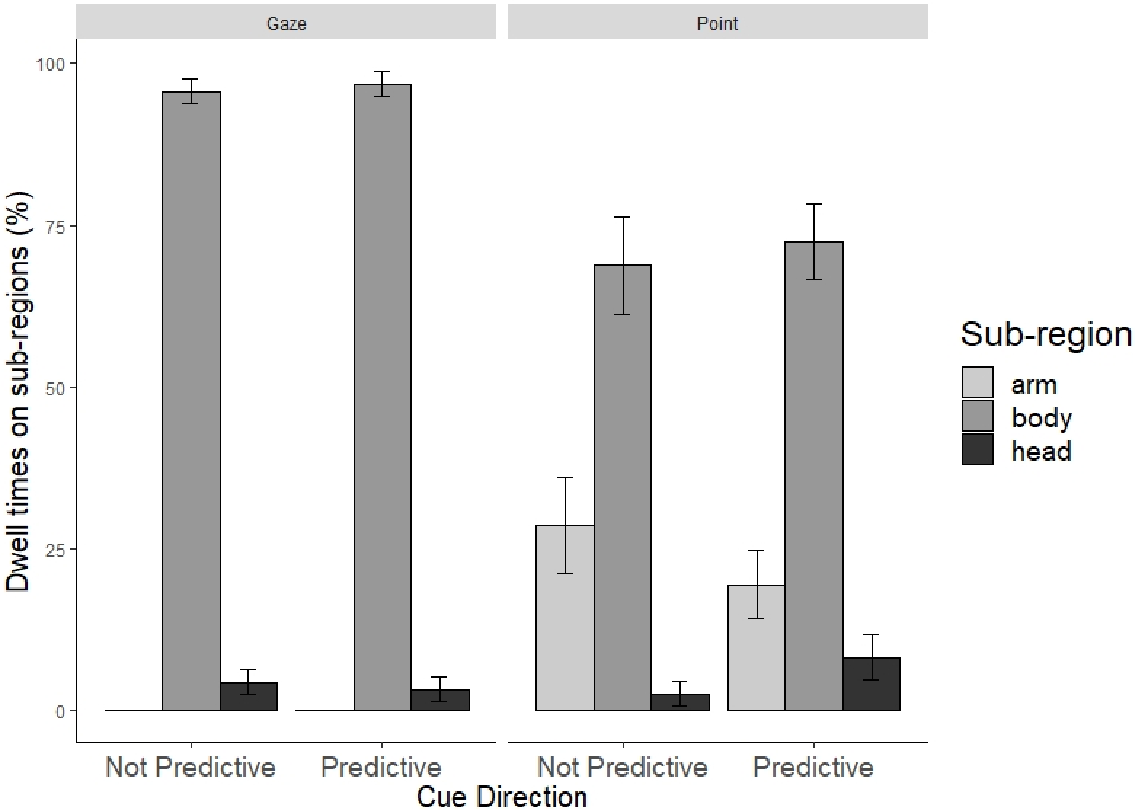

Fig. A.2 Dwell Times on the three body parts for the two social cues (gaze cue on the left, pointing cue on the right) and the two cue directions when actors' position was closer to the periphery of the scene.

A.5 Pictures of stimuli used in Experiment 5 and 6, Chapter 4.

Fig. A.3 Pictures of stimuli used in Experiment 5 (1st row) and 6 (2nd row), Chapter 4.